

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:17 ; Search time 12.3636 Seconds
(without alignments)
19.038 Million cell1 updates/sec

Title: B
Perfect score: 40
Sequence: 1 GSSFISPE 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 180334

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	11	4	US-09-608-810A-2
2	30	75.0	45	4	US-08-975-080-11
3	27	67.5	17	4	US-09-155-613A-46
4	27	67.5	35	2	US-08-724-194-11
5	27	67.5	35	4	US-09-171-482-7
6	26	65.0	11	2	US-08-934-222-42
7	26	65.0	11	2	US-08-933-402-42
8	26	65.0	11	2	US-09-207-621-42
9	26	65.0	11	2	US-08-532-818-42
10	26	65.0	11	3	US-09-231-797-42
11	26	65.0	11	3	US-08-934-224-42
12	26	65.0	11	3	US-08-933-843-42
13	26	65.0	11	4	US-08-934-223-42
14	26	65.0	11	4	US-09-113-492-42
15	25	62.5	19	1	US-07-977-444C-1
16	25	62.5	27	2	US-08-557-309B-46
17	25	62.5	27	3	US-08-834-306-46
18	25	62.5	27	4	US-08-993-674A-46
19	25	62.5	27	4	US-09-311-311C-8
20	25	62.5	27	4	US-09-256-976-46
21	25	62.5	31	4	US-08-706-344C-24
22	25	62.5	38	4	US-09-172-841-25
23	25	62.5	45	4	US-08-975-080-9
24	25	62.5	50	1	US-08-259-672-19
25	25	62.5	50	1	US-08-459-351-19
26	25	62.5	50	1	US-08-460-533-19
27	25	62.5	50	5	PCT-US94-0654-19

28	24	60.0	7	4	US-08-378-313-5	Sequence 5, Appl
29	24	60.0	12	1	US-08-196-989B-6	Sequence 6, Appl
30	24	60.0	12	1	US-08-760-936-6	Sequence 6, Appl
31	24	60.0	13	2	US-08-637-759B-129	Sequence 129, App
32	24	60.0	13	3	US-08-871-355A-129	Sequence 129, App
33	24	60.0	13	3	US-09-201-945-129	Sequence 129, App
34	24	60.0	15	2	US-08-553-257A-7	Sequence 3, Appl
35	24	60.0	15	4	US-09-268-480-3	Sequence 3, Appl
36	24	60.0	20	3	US-08-467-023-52	Sequence 24, Appl
37	24	60.0	20	5	PCT-US95-06726-24	Sequence 24, Appl
38	24	60.0	21	1	US-08-471-788C-70	Sequence 70, Appl
39	24	60.0	21	1	US-08-467-282B-70	Sequence 70, Appl
40	24	60.0	21	2	US-08-471-282A-70	Sequence 70, Appl
41	24	60.0	21	2	US-08-466-710C-70	Sequence 70, Appl
42	24	60.0	21	3	US-08-468-739C-70	Sequence 70, Appl
43	24	60.0	29	4	US-09-311-311C-15	Sequence 15, Appl
44	24	60.0	32	1	US-08-471-788C-21	Sequence 21, Appl
45	24	60.0	32	1	US-08-467-282B-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-608-810A-2
; Sequence 2, Application US/09608810A
; Patent No. 6420521
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Delisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: SCIP PEPTIDES
; FILE REFERENCE: 99-51
; CURRENT APPLICATION NUMBER: US/09/608,810A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/141,592
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-810A-2

Query Match 100.0%; Score 40; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSFISPE 8
DB 1 GSSFISPE 8

RESULT 2
US-08-975-080-11
; Sequence 11, Application US/08975080
; Patent No. 6245523
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,080
FILING DATE: 20-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,435
FILING DATE: 20-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-975-080-11

Query Match 75.0%; Score 30; DB 4; Length 45;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SFLSPE 8
Db 14 SFLSPE 19

RESULT 3

US-09-155-613A-46
Sequence 46, Application US/09155613A
Patent No. 6420120
GENERAL INFORMATION:
APPLICANT: Boulanger, Pierre
APPLICANT: Hong, Saw See
APPLICANT: Karayan, Lucie
TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
FILE REFERENCE: 032751-036
CURRENT APPLICATION NUMBER: US/09/155,613A
CURRENT FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: PCT/FR98/00184
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: FR 97/01005
PRIOR FILING DATE: 1997-01-30
PRIOR APPLICATION NUMBER: FR 97/11166
PRIOR FILING DATE: 1997-09-09
NUMBER OF SEQ ID NOS: 98
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Phagotopo
US-09-155-613A-46

Query Match 67.5%; Score 27; DB 4; Length 17;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 SSFLSPE 8
Db 1 NSFLDPE 7

RESULT 4

US-08-724-194-11

Sequence 11, Application US/08724194
Patent No. 5824875
GENERAL INFORMATION:
APPLICANT: RANU, RAJINDER S.
TITLE OF INVENTION: ONE-AMINOCYCLOPROPANE-1-CARBOXYLATE
TITLE OF INVENTION: SYNTHASE GENES FROM PELARGONIUM TO CONTROL ETHYLENE LEVELS
TITLE OF INVENTION: IN GERANIUMS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SANTANGELO LAW OFFICES PC
STREET: 315 WEST OAK STREET, STE 701
CITY: FORT COLLINS
STATE: CO
COUNTRY: USA
ZIP: 80521
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,194
FILING DATE: 01-OCT-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SANTANGELO, LUKE
REGISTRATION NUMBER: 31,997
TELECOMMUNICATION INFORMATION:
TELEPHONE: (970) 224-3100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-724-194-11

Query Match 67.5%; Score 27; DB 2; Length 35;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8
Db 5 GSSFLCSE 12

RESULT 5

US-09-171-482-7
Sequence 7, Application US/09171482A
Patent No. 6184449
GENERAL INFORMATION:
APPLICANT: Ranu, Rajinder S.
TITLE OF INVENTION: A 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE GENE FROM
TITLE OF INVENTION: ROSA TO CONTROL ETHYLENE LEVELS IN ROSES
FILE REFERENCE: TAGAWA-ROSE
CURRENT APPLICATION NUMBER: US/09/171,482A
CURRENT FILING DATE: 1998-10-19
EARLIER APPLICATION NUMBER: PCT/US97/17644, Published under WO98/14465; US5,824,875
EARLIER FILING DATE: 1997-Sept-30, Published 1998-April-09; 1996-Oct-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Word Perfect 6.1
SEQ ID NO 7
LENGTH: 35
TYPE: PRT
ORGANISM: Rosa kardinal
US-09-171-482-7

Query Match 67.5%; Score 27; DB 4; Length 35;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8

Db 5 GSSFLCSB 12

RESULT 6

US-08-934-222-42

Sequence 42, Application US/08934222

Patent No. 5928896

GENERAL INFORMATION:

APPLICANT: EVANS, Herbert J.

APPLICANT: KINI, R. Manjunatha

TITLE OF INVENTION: Polypeptides That Include Conformation-

TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction

TITLE OF INVENTION: Site

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner

STREET: Suite 500, 3000 K Street NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20007

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/934,222

FILING DATE: 19-SEPT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/532,818

FILING DATE: 03-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/143,364

FILING DATE: 29-OCT-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/051,741

FILING DATE: 23-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Isaacson, John P.

REGISTRATION NUMBER: 33,751

REFERENCE/DOCKET NUMBER: 040433/0148

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-934-222-42

Query Match

Best Local Similarity 65.0%; Score 26; DB 2; Length 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 SSFLSPE 8

Db 5 SSTLSPE 11

RESULT 7

US-08-933-402-42

Sequence 42, Application US/08933402

Patent No. 5948887

GENERAL INFORMATION:

APPLICANT: EVANS, Herbert J.

APPLICANT: KINI, R. Manjunatha

TITLE OF INVENTION: Polypeptides That Include Conformation-

TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction

TITLE OF INVENTION: Site

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner

STREET: Suite 500, 3000 K Street NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20007

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/207,621

FILING DATE: 03-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/532,818

FILING DATE: 21-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/143,364

STREET: Suite 500, 3000 K Street NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20007

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,402

FILING DATE: 19-SEPT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/532,818

FILING DATE: 03-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/143,364

FILING DATE: 29-OCT-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/051,741

FILING DATE: 23-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Isaacson, John P.

REGISTRATION NUMBER: 33,751

REFERENCE/DOCKET NUMBER: 040433/0148

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-933-402-42

Query Match

Best Local Similarity 65.0%; Score 26; DB 2; Length 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 SSFLSPE 8

Db 5 SSTLSPE 11

RESULT 8

US-09-207-621-42

Sequence 42, Application US/09207621

Patent No. 5952465

GENERAL INFORMATION:

APPLICANT: EVANS, Herbert J.

APPLICANT: KINI, R. Manjunatha

TITLE OF INVENTION: Polypeptides That Include Conformation-

TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction

TITLE OF INVENTION: Site

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner

STREET: Suite 500, 3000 K Street NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20007

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/207,621

FILING DATE: 03-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/532,818

FILING DATE: 21-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/143,364

STREET: Suite 500, 3000 K Street NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20007

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/207,621

FILING DATE: 03-MAY-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/532,818

FILING DATE: 21-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/143,364

STREET: Suite 500, 3000 K Street NW

/ FILING DATE: 29-OCT-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: U.S. 08/051,741
/ FILING DATE: 23-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Isaacson, John P.
/ REGISTRATION NUMBER: 33,751
/ REFERENCE/DOCKET NUMBER: 040433/0148
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
US-09-207-621-42

Query Match 65.0%; Score 26; DB 2; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 9
US-08-532-818-42
/ Sequence 42, Application US/08532818
/ Patent No. 5965698
/ GENERAL INFORMATION:
/ APPLICANT: EVANS, Herbert J.
/ APPLICANT: KINI, R. Manjunatha
/ TITLE OF INVENTION: Polypeptides That Include Conformation-
/ TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
/ NUMBER OF SEQUENCES: 153
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: Suite 500, 3000 K Street NW
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20007
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 03-MAY-1996
/ PRIOR APPLICATION NUMBER: PCT/US94/04294
/ FILING DATE: 21-APR-1994
/ APPLICATION NUMBER: U.S. 08/051,741
/ FILING DATE: 29-OCT-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: U.S. 08/143,364
/ FILING DATE: 23-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Isaacson, John P.
/ REGISTRATION NUMBER: 33,751
/ REFERENCE/DOCKET NUMBER: 040433/0148
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
US-08-532-818-42

Query Match 65.0%; Score 26; DB 2; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 9
US-08-532-818-42
/ Sequence 42, Application US/08532818
/ Patent No. 5965698
/ GENERAL INFORMATION:
/ APPLICANT: EVANS, Herbert J.
/ APPLICANT: KINI, R. Manjunatha
/ TITLE OF INVENTION: Polypeptides That Include Conformation-
/ TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
/ NUMBER OF SEQUENCES: 153
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: Suite 500, 3000 K Street NW
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20007
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 03-MAY-1996
/ PRIOR APPLICATION NUMBER: PCT/US94/04294
/ FILING DATE: 21-APR-1994
/ APPLICATION NUMBER: U.S. 08/051,741
/ FILING DATE: 29-OCT-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: U.S. 08/143,364
/ FILING DATE: 23-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Isaacson, John P.
/ REGISTRATION NUMBER: 33,751
/ REFERENCE/DOCKET NUMBER: 040433/0148
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
US-08-532-818-42

Query Match 65.0%; Score 26; DB 2; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 10
US-09-231-797-42
/ Sequence 42, Application US/09231797
/ Patent No. 6084066
/ GENERAL INFORMATION:
/ APPLICANT: EVANS, Herbert J.
/ APPLICANT: KINI, R. Manjunatha
/ TITLE OF INVENTION: Polypeptides That Include Conformation-
/ TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
/ NUMBER OF SEQUENCES: 153
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: Suite 500, 3000 K Street NW
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20007
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 03-MAY-1996
/ PRIOR APPLICATION NUMBER: PCT/US94/04294
/ FILING DATE: 21-APR-1994
/ APPLICATION NUMBER: U.S. 08/143,364
/ FILING DATE: 29-OCT-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: U.S. 08/051,741
/ FILING DATE: 23-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Isaacson, John P.
/ REGISTRATION NUMBER: 33,751
/ REFERENCE/DOCKET NUMBER: 040433/0148
/ INFORMATION FOR SEQ ID NO: 42:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
US-09-231-797-42

Query Match 65.0%; Score 26; DB 3; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 11
US-08-934-224-42
/ Sequence 42, Application US/08934224
/ Patent No. 6100044
/ GENERAL INFORMATION:
/ APPLICANT: EVANS, Herbert J.
/ APPLICANT: KINI, R. Manjunatha
/ TITLE OF INVENTION: Polypeptides That Include Conformation-
/ TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
/ NUMBER OF SEQUENCES: 153

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: Suite 500, 3000 K Street NW
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20007
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/934,224
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/532,818
;; FILING DATE: 03-MAY-1996
;; APPLICATION NUMBER: PCT/US94/04294
;; FILING DATE: 21-APR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/143,364
;; FILING DATE: 29-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/051,741
;; FILING DATE: 23-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Isaacson, John P.
;; REGISTRATION NUMBER: 33,751
;; REFERENCE/DOCKET NUMBER: 040433/0148
;; INFORMATION FOR SEQ ID NO: 42:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
;; US-08-934-224-42
;;
Query Match 65.0%; Score 26; DB 3; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;;
QY 2 SSFLSPE 8
 |||
Db 5 SSFLSPE 11
;;
RESULT 12
US-08-933-843-42
;; Sequence 42, Application US/08933843
;; Patent No. 6111069
;; GENERAL INFORMATION:
;; APPLICANT: EVANS, Herbert J.
;; APPLICANT: KINI, R. Manjunatha
;; TITLE OF INVENTION: Polypeptides That Include Conformation-
;; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
;; NUMBER OF SEQUENCES: 153
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: Suite 500, 3000 K Street NW
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20007
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/933,843
;; FILING DATE: 19-SEPT-1997
;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 08/532,818
;; FILING DATE: 03-MAY-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/143,364
;; FILING DATE: 29-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/051,741
;; FILING DATE: 23-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Isaacson, John P.
;; REGISTRATION NUMBER: 33,751
;; REFERENCE/DOCKET NUMBER: 040433/0148
;; INFORMATION FOR SEQ ID NO: 42:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
;; US-08-933-843-42
;;
Query Match 65.0%; Score 26; DB 3; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;;
QY 2 SSFLSPE 8
 |||
Db 5 SSFLSPE 11
;;
RESULT 13
US-08-934-223-42
;; Sequence 42, Application US/08934223
;; Patent No. 6147189
;; GENERAL INFORMATION:
;; APPLICANT: EVANS, Herbert J.
;; APPLICANT: KINI, R. Manjunatha
;; TITLE OF INVENTION: Polypeptides That Include Conformation-
;; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
;; NUMBER OF SEQUENCES: 153
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: Suite 500, 3000 K Street NW
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20007
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/934,223
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/532,818
;; FILING DATE: 03-MAY-1996
;; APPLICATION NUMBER: PCT/US94/04294
;; FILING DATE: 21-APR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/143,364
;; FILING DATE: 29-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/051,741
;; FILING DATE: 23-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Isaacson, John P.
;; REGISTRATION NUMBER: 33,751
;; REFERENCE/DOCKET NUMBER: 040433/0148
;; INFORMATION FOR SEQ ID NO: 42:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid

```

; TOPOLOGY: linear
US-08-934-223-42
Query Match 65.0%; Score 26; DB 4; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 14
US-09-413-492-42
; Sequence 42, Application US/09413492
; Patent No. 6258550
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; TITLE OF INVENTION: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; TITLE OF INVENTION: Site
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/413,492
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-413-492-42
Query Match 65.0%; Score 26; DB 4; Length 11;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFLSPE 8
Db 5 SSTLSPE 11

RESULT 15
US-07-977-444C-1
; Sequence 1, Application US/07977444C
; Patent No. 5449750

```

```

; GENERAL INFORMATION:
; APPLICANT: K. KIMURA, et al.
; TITLE OF INVENTION: NOVEL PROLYL ENDOPEPTIDASE
; TITLE OF INVENTION: INHIBITORS SNA-115 AND SNA-115T AND
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION AND
; TITLE OF INVENTION: PRODUCTIVE STRAIN THEREOF
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5 1/4 inch diskette
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,444C
; FILING DATE: 17 - FEBRUARY 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wayne, Milton J.
; REGISTRATION NUMBER: 17,906
; REFERENCE/DOCKET NUMBER: U-Wp-4947
; TELEPHONE: 212-683-8150
; TELEFAX: 212-532-4285
; TELEX: 423794
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19
; TYPE: AMINO ACID
; TOPOLOGY: CYCLIC or LINEAR
US-07-977-444C-1
Query Match 62.5%; Score 25; DB 1; Length 19;
Best Local Similarity 57.1%; Pred. No. 69;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSFLSP 7
Db 13 GHTFISP 19

Search completed: January 10, 2003, 15:59:54
Job time : 12.3636 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 10, 2003, 15:55:17 ; Search time 6.18182 Seconds
(without alignments)
19.038 Million cell updates/sec

Title: A
Perfect score: 20
Sequence: 1 gseq 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 180334

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	11	4	US-09-608-810A-2
2	20	100.0	13	2	US-08-637-759B-129
3	20	100.0	13	3	US-08-871-355A-129
4	20	100.0	13	4	US-09-201-945-129
5	20	100.0	17	2	US-08-960-128-1
6	20	100.0	20	5	PCT-US95-06726-24
7	20	100.0	26	1	US-07-942-245-383
8	20	100.0	35	2	US-08-724-194-11
9	20	100.0	35	4	US-09-171-482-7
10	17	85.0	4	6	5198359-8
11	17	85.0	4	6	5284931-17
12	17	85.0	4	6	5449756-11
13	17	85.0	5	2	US-08-618-696-13
14	17	85.0	5	3	US-09-033-753-13
15	17	85.0	5	6	5438119-11
16	17	85.0	7	3	US-08-889-841B-54
17	17	85.0	7	3	US-08-889-841B-56
18	17	85.0	8	2	US-08-276-967-7
19	17	85.0	8	4	US-08-444-818-642
20	17	85.0	8	4	US-08-444-818-643
21	17	85.0	8	4	US-08-444-818-644
22	17	85.0	8	4	US-08-444-818-645
23	17	85.0	8	4	US-08-444-818-646
24	17	85.0	8	4	US-09-043-731-7
25	17	85.0	9	2	US-08-350-260A-435
26	17	85.0	9	3	US-08-925-002-52
27	17	85.0	10	1	US-07-670-296-9

28	17	85.0	10	1	US-08-093-781-10	Sequence 10, Appl
29	17	85.0	10	1	US-08-416-950-3	Sequence 3, Appl
30	17	85.0	10	2	US-08-618-696-9	Sequence 9, Appl
31	17	85.0	10	2	US-08-361-517-19	Sequence 19, Appl
32	17	85.0	10	2	US-08-469-830-3	Sequence 3, Appl
33	17	85.0	10	2	US-09-033-753-9	Sequence 9, Appl
34	17	85.0	10	5	PCT-US93-07964-19	Sequence 19, Appl
35	17	85.0	11	2	US-08-618-696-5	Sequence 5, Appl
36	17	85.0	11	2	US-08-618-696-18	Sequence 18, Appl
37	17	85.0	11	3	US-09-033-753-5	Sequence 5, Appl
38	17	85.0	11	3	US-09-033-753-18	Sequence 18, Appl
39	17	85.0	13	1	US-08-189-772-2	Sequence 2, Appl
40	17	85.0	13	1	US-08-082-849B-17	Sequence 17, Appl
41	17	85.0	13	1	US-08-188-277B-11	Sequence 11, Appl
42	17	85.0	13	5	PCT-US94-01624-17	Sequence 17, Appl
43	17	85.0	14	1	US-08-188-277B-24	Sequence 24, Appl
44	17	85.0	15	1	US-07-609-716-16	Sequence 16, Appl
45	17	85.0	15	1	US-07-609-716-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-09-608-810A-2
; Sequence 2, Application US/09608810A
; Patent No. 6420521
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Jasper, Stephen R.
; APPLICANT: Delsher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: SGIP PEPTIDES
; FILE REFERENCE: 99-51
; CURRENT APPLICATION NUMBER: US/09/608, 810A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/141,592
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-810A-2

Query Match 100.0%; Score 20; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
Db 1 GSSF 4

RESULT 2
US-08-637-759B-129
; Sequence 129, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS: 501
; ADDRESS: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: 1201 West Peachtree Street
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

US-02-062002 in last

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 129:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-129

Query Match 100.0%; Score 20; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
Db 5 GSSF 8

RESULT 3
US-08-871-355A-129
Sequence 129, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795

INFORMATION FOR SEQ ID NO: 129:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-129

Query Match 100.0%; Score 20; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
Db 5 GSSF 8

RESULT 4
US-09-201-945-129
Sequence 129, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 129:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-129

Query Match 100.0%; Score 20; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
Db 5 GSSF 8

RESULT 5
US-08-960-128-1
Sequence 1, Application US/08960128
Patent No. 5951985
GENERAL INFORMATION:
APPLICANT: Butler, Sandra M.
APPLICANT: Pomato, Nicholas
APPLICANT: Bog, Ebo
APPLICANT: Hanna, Michael G.
APPLICANT: Haspel, Martin V.
APPLICANT: Hoover, Herbert C.
TITLE OF INVENTION: Tumor Associated Epitopes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5951985el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,591
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
TELEFAX: (301) 977-0847
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-960-128-1

Query Match 100.0%; Score 20; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||||
DB 4 GSSF 7

RESULT 6
PCT-US95-06726-24
Sequence 24, Application PC/TUS9506726
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Ligands for Induction of Antigen Specific Apoptosis in
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06726
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/253,783
FILING DATE: 03 JUNE 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-016PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-06726-24

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||||
DB 9 GSSF 12

RESULT 7
US-07-942-245-383
Sequence 383, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: REES, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mfon, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 383:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-383

Query Match 100.0%; Score 20; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 8 GSSF 11

RESULT 8

US-08-724-194-11
; Sequence 11, Application US/08724194
; Patent No. 5824875
; GENERAL INFORMATION:
; APPLICANT: RANU, RAJINDER S.
; TITLE OF INVENTION: ONE-AMINOCYCLOPROPANE-1-CARBOXYLATE
; TITLE OF INVENTION: SYNTHASE GENES FROM PELARGONIUM TO CONTROL ETHYLENE LEVELS
; TITLE OF INVENTION: IN GERANIUMS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SANTIANGELO LAW OFFICES PC
; STREET: 315 WEST OAK STREET, STE 701
; CITY: FORT COLLINS
; STATE: CO
; COUNTRY: USA
; ZIP: 80521
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,194
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SANTIANGELO, LUKE
; REGISTRATION NUMBER: 31,997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 224-3100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-724-194-11

Query Match 100.0%; Score 20; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 5 GSSF 8

RESULT 9

US-09-171-482-7
; Sequence 7, Application US/09171482A
; Patent No. 6184449
; GENERAL INFORMATION:
; APPLICANT: Ranu, Rajinder S.
; TITLE OF INVENTION: A 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE GENE FROM
; TITLE OF INVENTION: ROSA TO CONTROL ETHYLENE LEVELS IN ROSES
; FILE REFERENCE: TAGAWA-ROSE
; CURRENT APPLICATION NUMBER: US/09/171,482A

; CURRENT FILING DATE: 1998-10-19
; EARLIER APPLICATION NUMBER: PCT/US97/17644, Published under WO98/14465; US5,824,875
; EARLIER FILING DATE: 1997-Sept-30, Published 1998-April-09; 1996-Oct-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Word Perfect 6.1
; SEQ ID NO 7
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Rosa kardinal
US-09-171-482-7

Query Match 100.0%; Score 20; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 5 GSSF 8

RESULT 10

5198359-8
; Patent No. 5198359
; APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI;
; MINAMOTO, SEIJO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI;
; TSUDO, MITSURU; KARASUYMA, HAJIME
; TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,059
; FILING DATE: 05-MAR-1990
; SEQ ID NO: 8
; LENGTH: 4
5198359-8

Query Match 85.0%; Score 17; DB 6; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 1 GASF 4

RESULT 11

5284931-17
; Patent No. 5284931
; APPLICANT: SPRINGER, TIMOTHY A.; ROTHLEIN, ROBERT; MARLIN,
; STEVEN D.; DUSTIN, MICHAEL L.
; TITLE OF INVENTION: INTERCELLULAR ADHESION MOLECULES AND
; THEIR BINDING LIGANDS
; NUMBER OF SEQUENCES: 41
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/515,478
; FILING DATE: 27-APR-1990
; SEQ ID NO: 17
; LENGTH: 4
5284931-17

Query Match 85.0%; Score 17; DB 6; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 1 GASF 4

RESULT 12

5449756-11
; Patent No. 5449756
; APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI; MINAMOTO,
; SEIJO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI; TSUDO,

;/MITSURU,KARASUYAMA, HAJIME
/ TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
/ NUMBER OF SEQUENCES: 12
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/88,592
/ FILING DATE: 9-JUL-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 865,155
/ FILING DATE: 08-APR-1992
/ APPLICATION NUMBER: 487,059
/ FILING DATE: 05-MAR-1990
/ SEQ ID NO:11:
/ LENGTH: 4
5449756-11

Query Match 85.0%; Score 17; DB 6; Length 4;
Best Local Similarity 75.0%; Pred.No.1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 1 GSAF 4

RESULT 13
US-08-618-696-13
/ Sequence 13, Application US/08618696
/ Patent No. 5861475
/ GENERAL INFORMATION:
/ APPLICANT: COOPER, Jr., J. ALLEN D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: ARNOLD, WHITE & DURKEE
/ STREET: P. O. BOX 4433
/ CITY: HOUSTON
/ STATE: TEXAS
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/618,696
/ FILING DATE: 20-MAR-1996
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/995,269
/ FILING DATE: 12/21/92
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PARKER, DAVID L.
/ REGISTRATION NUMBER: 32,165
/ REFERENCE/DOCKET NUMBER: UOAB:002/PAR
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512-320-7200
/ TELEFAX: 512-474-7577
/ TELEX: NOT APPLICABLE
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acid residues
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-618-696-13

Query Match 85.0%; Score 17; DB 2; Length 5;
Best Local Similarity 75.0%; Pred.No.1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 1 GSAF 4

Db 1 GSAF 4

RESULT 14
US-09-033-753-13
/ Sequence 13, Application US/09033753
/ Patent No. 6017883
/ GENERAL INFORMATION:
/ APPLICANT: COOPER, Jr., J. ALLEN D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: ARNOLD, WHITE & DURKEE
/ STREET: P. O. BOX 4433
/ CITY: HOUSTON
/ STATE: TEXAS
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/033,753
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/618,696
/ FILING DATE: 20-MAR-1996
/ APPLICATION NUMBER: 07/995,269
/ FILING DATE: 12/21/92
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PARKER, DAVID L.
/ REGISTRATION NUMBER: 32,165
/ REFERENCE/DOCKET NUMBER: UOAB:002/PAR
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 512-320-7200
/ TELEFAX: 512-474-7577
/ TELEX: NOT APPLICABLE
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acid residues
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-09-033-753-13

Query Match 85.0%; Score 17; DB 3; Length 5;
Best Local Similarity 75.0%; Pred.No.1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 1 GSAF 4

RESULT 15
5438119-11
/ Patent No. 5438119
/ APPLICANT: RUTER, William,Santi, Daniel
/ TITLE OF INVENTION: METHOD OF OBTAINING A PEPTIDE WITH DESIRED
/ TARGET PROPERTY
/ NUMBER OF SEQUENCES: 16
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/981,759
/ FILING DATE: 25-NOV-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 525,899
/ FILING DATE: 18-MAY-1990
/ APPLICATION NUMBER: 189,318

Mon Jan 13 09:26:14 2003

a.ra1

Page 6

; FILING DATE: 02-MAY-1988
;SEQ ID NO:11:
; LENGTH: 5
5438119-11

Query Match 85.0%; Score 17; DB 6; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.9e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GSSF 4
Db 1 GSAF 4

Search completed: January 10, 2003, 15:59:54
Job time : 8.18182 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:17 ; Search time 15.4545 Seconds
(without alignments)
19.038 Million cell updates/sec

Title: C
Perfect score: 50
Sequence: 1 gsfaklqpr 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 180334

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/prodata/1/1aa/PCITUS_COMB.pep: *
6: /cgn2_6/prodata/1/1aa/backfill1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	58.0	21	4	US-08-875-277A-13	Sequence 13, Appl
2	28	56.0	17	4	US-08-836-922-18	Sequence 18, Appl
3	28	56.0	32	6	5182210-22	Patent No. 5182210
4	28	56.0	50	6	5182210-16	Patent No. 5182210
5	27	54.0	9	3	US-08-159-339A-824	Sequence 824, App
6	27	54.0	10	3	US-08-159-339A-854	Sequence 854, App
7	27	54.0	18	4	US-09-385-740B-17	Sequence 17, Appl
8	27	54.0	19	4	US-09-001-984C-72	Sequence 72, Appl
9	27	54.0	19	4	US-09-441-502B-54	Sequence 54, Appl
10	27	54.0	20	4	US-09-385-740B-18	Sequence 18, Appl
11	27	54.0	21	4	US-09-385-740B-19	Sequence 19, Appl
12	27	54.0	23	2	US-08-548-974-24	Sequence 24, Appl
13	27	54.0	25	2	US-08-548-974-15	Sequence 15, Appl
14	27	54.0	27	1	US-07-919-731-2	Sequence 2, Appl
15	27	54.0	27	1	US-08-287-957-107	Sequence 107, Appl
16	27	54.0	27	2	US-08-337-127-2	Sequence 2, Appl
17	27	54.0	27	2	US-08-548-974-17	Sequence 17, Appl
18	27	54.0	27	3	US-08-874-775-46	Sequence 46, Appl
19	27	54.0	27	3	US-09-260-846-2	Sequence 2, Appl
20	27	54.0	27	6	5460801-2	Patent No. 5460801
21	26	52.0	15	1	US-08-434-705B-18	Sequence 18, Appl
22	26	52.0	15	2	US-09-086-201-18	Sequence 18, Appl
23	26	52.0	31	1	US-08-296-644-1	Sequence 1, Appl
24	26	52.0	31	2	US-08-578-240-1	Sequence 1, Appl
25	26	52.0	31	3	US-08-950-925-6	Sequence 6, Appl
26	26	52.0	37	4	US-08-905-223-321	Sequence 321, Appl
27	26	52.0	39	1	US-08-342-101-1	Sequence 1, Appl

28	26	52.0	39	2	US-08-535-361-1	Sequence 1, Appl
29	26	52.0	39	5	PCT-US94-05468-1	Sequence 1, Appl
30	25	50.0	10	2	US-08-556-597-171	Sequence 171, App
31	25	50.0	13	2	US-08-637-759B-129	Sequence 129, App
32	25	50.0	13	3	US-08-871-355A-129	Sequence 129, App
33	25	50.0	13	4	US-09-201-945-129	Sequence 129, App
34	25	50.0	12	2	US-08-143-124-24	Sequence 145, App
35	24	48.0	29	4	US-09-143-124-24	Sequence 145, App
36	24	48.0	20	2	US-08-763-374-2	Sequence 13, Appl
37	24	48.0	21	4	US-09-028-937-13	Sequence 119, App
38	24	48.0	23	4	US-09-220-528-119	Sequence 12, Appl
39	24	48.0	28	1	US-08-340-428B-12	Sequence 12, Appl
40	24	48.0	28	5	PCT-US93-07306-12	Sequence 24, Appl
41	24	48.0	29	4	US-09-082-358B-34	Sequence 34, Appl
42	24	48.0	34	1	US-08-155-171B-8	Sequence 8, Appl
43	24	48.0	34	2	US-08-435-998-8	Sequence 73, Appl
44	24	48.0	34	3	US-08-486-099-73	Sequence 74, Appl
45	24	48.0	34	3	US-08-486-099-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-08-875-277A-13
; Sequence 13, Application US/08875277A
; Patent No. 6171808
; GENERAL INFORMATION:
; APPLICANT: SOIRRELL, DAVID J.
; APPLICANT: LOWE, CHRISTOPHER R.
; APPLICANT: WHITE, PETER J.
; APPLICANT: MURRAY, JAMES A.H.
; TITLE OF INVENTION: MUTANT LUCIFERASES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,277A
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9501172.2
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9508301.0
; FILING DATE: 24-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET INFORMATION: 124-588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-875-277A-13
Query Match 58.0%; Score 29; DB 4; Length 21;
Best Local Similarity 60.0%; Pred. No. 13;

```
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSSFAKLQPR 10
Db 8 GSSRVDLPK 17

RESULT 2
US-08-836-922-18
; Sequence 18, Application US/08836922
; Patent No. 6159711
; GENERAL INFORMATION:
; APPLICANT: INNES PROUDFOOT, AMANDA ELIZABETH
; APPLICANT: WELLS, TIMOTHY NIGEL CARL
; TITLE OF INVENTION: RANTES PEPTIDE AND FRAGMENTS AND
; TITLE OF INVENTION: COMPOSITIONS COMPRISING IT FOR TREATMENT OF INFLAMMATION
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,922
; FILING DATE: 23-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9424835.8
; FILING DATE: 08-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9512319.6
; FILING DATE: 16-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1430-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4011
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-836-922-18

Query Match 56.0%; Score 28; DB 4; Length 17;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSSFAKLQPR 10
Db 4 GSSRVDLPK 13

RESULT 3
5182210-22
; Patent No. 5182210
; APPLICANT: BINNS, MATTHEW M.;BOURNSELL, MICHAEL E.G.;
; CAMPBELL, JOAN I.A.;TOMLEY, FIONA M.
; TITLE OF INVENTION: FOWLPOX VIRUS PROMOTERS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,608
; FILING DATE: 21-OCT-1988
```

```
;SEQ ID NO:22:
; LENGTH: 32
5182210-22

Query Match 56.0%; Score 28; DB 6; Length 32;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSSFAKLQPR 10
Db 5 GSSRVDLPK 14

RESULT 4
5182210-16
; Patent No. 5182210
; APPLICANT: BINNS, MATTHEW M.;BOURNSELL, MICHAEL E.G.;
; CAMPBELL, JOAN I.A.;TOMLEY, FIONA M.
; TITLE OF INVENTION: FOWLPOX VIRUS PROMOTERS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,608
; FILING DATE: 21-OCT-1988
;SEQ ID NO:16:
; LENGTH: 50
5182210-16

Query Match 56.0%; Score 28; DB 6; Length 50;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSSFAKLQPR 10
Db 23 GSSRVDLPK 32

RESULT 5
US-08-159-339A-824
; Sequence 824, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
```

REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 824:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-159-339A-824

Query Match 54.0%; Score 27; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSFAXLQPR 10
Db 1 SAFAGLGPR 9

RESULT 6
US-08-159-339A-854
Sequence 854, Application US/08159339A
Patent No. 6037135
GENERAL INFORMATION:
APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esben
TITLE OF INVENTION: HLA Binding peptides and their
NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 854:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-159-339A-854

Query Match 54.0%; Score 27; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSFAXLQPR 10
Db 2 SAFAGLGPR 10

RESULT 7
US-09-385-740B-17
Sequence 17, Application US/09385740B
Patent No. 6348320
GENERAL INFORMATION:
APPLICANT: Eyre, David
TITLE OF INVENTION: CARTILAGE RESORPTION ASSAYS
FILE REFERENCE: WROS-1-14269
CURRENT APPLICATION NUMBER: US/09/385,740B
CURRENT FILING DATE: 1999-08-30
PRIOR APPLICATION NUMBER: US 60/142,274
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: US 60/141,574
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 09/335,098
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: US 60/089,823
PRIOR FILING DATE: 1998-06-19
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)-(18)
OTHER INFORMATION: syn- corr. to C-terminal telopeptide seq. of hu. type II collagen
US-09-385-740B-17

Query Match 54.0%; Score 27; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSFAXLQPR 10
Db 4 SAFAGLGPR 12

RESULT 8
US-09-001-984C-72
Sequence 72, Application US/09001984C
Patent No. 6245331
GENERAL INFORMATION:
APPLICANT: Laai, Suman
APPLICANT: Zolla-Pazner, Susan
APPLICANT: Bellisle, John T.
TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
FILE REFERENCE: NYU-011
CURRENT APPLICATION NUMBER: US/09/001,984C
CURRENT FILING DATE: 1997-12-31
PRIOR APPLICATION NUMBER: 60/034,003
PRIOR FILING DATE: 1996-12-31
NUMBER OF SEQ ID NOS: 106
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 72
LENGTH: 19
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-001-984C-72

```
Query Match          54.0%; Score 27; DB 4; Length 19;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 SFAKLQPR 10
Db 12 SFAVLEPK 19

RESULT 9
US-09-441-502B-54
; Sequence 54, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-54

Query Match          54.0%; Score 27; DB 4; Length 19;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GSSFAKLQPR 9
Db 2 GSGFPETQP 10

RESULT 10
US-09-385-740B-18
; Sequence 18, Application US/09385740B
; Patent No. 6348320
; GENERAL INFORMATION:
; APPLICANT: Eyre, David
; TITLE OF INVENTION: CARTILAGE RESORPTION ASSAYS
; FILE REFERENCE: WROS-1-14269
; CURRENT APPLICATION NUMBER: US/09/385,740B
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 60/142,274
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: US 60/141,574
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/335,098
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/089,823
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: syn- corr. to C-terminal telopeptide seq. of hu. type II collagen
US-09-385-740B-18

Query Match          54.0%; Score 27; DB 4; Length 20;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSSFAKLQPR 10
Db 4 SAFAGLQPR 12

US-09-385-740B-19
; Sequence 19, Application US/09385740B
; Patent No. 6348320
; GENERAL INFORMATION:
; APPLICANT: Eyre, David
; TITLE OF INVENTION: CARTILAGE RESORPTION ASSAYS
; FILE REFERENCE: WROS-1-14269
; CURRENT APPLICATION NUMBER: US/09/385,740B
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 60/142,274
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: US 60/141,574
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/335,098
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/089,823
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: syn- corr. to C-terminal telopeptide seq. of hu. type II collagen
US-09-385-740B-19

Query Match          54.0%; Score 27; DB 4; Length 21;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSSFAKLQPR 10
Db 4 SAFAGLQPR 12

RESULT 12
US-08-548-974-24
; Sequence 24, Application US/08548974
; Patent No. 5939529
; GENERAL INFORMATION:
; APPLICANT: Potempa, Lawrence A
; TITLE OF INVENTION: Methods And Kits For Stimulating
; TITLE OF INVENTION: Production Of Megakaryocytes And Thrombocytes
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS HOFER GILSON & LIONE
; STREET: P.O. Box 10395
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60610
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/548,974
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/202,033
; FILING DATE: 23-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 2545/90
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 321-4282
TELEFAX: (312) 321-4299
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-06-548-974-24

Query Match 54.0%; Score 27; DB 2; Length 23;
Best Local Similarity 40.0%; Pred. No. 37;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 10
|:|:::|
Db 2 GTVFSRMPPR 11

RESULT 13
US-06-548-974-15
Sequence 15, Application US/08548974
Patent No. 5939529
GENERAL INFORMATION:
APPLICANT: Potempa, Lawrence A
TITLE OF INVENTION: Methods And Kits For Stimulating
TITLE OF INVENTION: Production Of Megakaryocytes And Thrombocytes
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS HOFER GILSON & LYONE
STREET: P.O. Box 10395
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/548,974
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/202,033
FILING DATE: 23-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 2545/90
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 321-4282
TELEFAX: (312) 321-4299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-06-548-974-15

Query Match 54.0%; Score 27; DB 2; Length 25;
Best Local Similarity 40.0%; Pred. No. 41;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 10
|:|:::|
Db 3 GTVFSRMPPR 12

RESULT 14

US-07-919-731-2
Sequence 2, Application US/07919731
Patent No. 5439884
GENERAL INFORMATION:
APPLICANT: Spindel, Eliot R.
APPLICANT: Nagalla, Srinivasa R.
APPLICANT: Vijayaraghavan, Srinivasan
APPLICANT: Archibond, Anthony
TITLE OF INVENTION: METHOD OF CONTROLLING FERTILIZATION
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/919,731
FILING DATE: 19920727
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T. and Tsao, Y. Rocky
REGISTRATION NUMBER: 30,162 and 34,053
REFERENCE/DOCKET NUMBER: 00537/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
STRANDEDNESS: N/A
TOPOLOGY: N/A
FEATURES:
OTHER INFORMATION: The sequence has an amide C-terminus (i.e., CO₂NH₂), rather th

US-07-919-731-2

Query Match 54.0%; Score 27; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 10
|:|:::|
Db 8 GTVLAKYPR 17

RESULT 15
US-08-287-957-107
Sequence 107, Application US/08287957
Patent No. 5552520
GENERAL INFORMATION:
APPLICANT: HYUK KIM SUN
APPLICANT: RILEY KEYES, SUSAN
APPLICANT: MOREAU, SYLVIANE
APPLICANT: XIN DONG, ZHENG
APPLICANT: TAYLOR, JOHN
TITLE OF INVENTION: THERAPEUTIC PEPTIDE DERIVATIVES
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/287.957
FILING DATE: 09-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00537/100001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-287-957-107

Query Match 54.0%; Score 27; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GSSFAKIQPR 10
DB 8 GTVLAKMYP 17

Search completed: January 10, 2003, 15:59:55
Job time : 16.4545 secs

Gly-Ser-Ser-Phe

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:17 ; Search time 15.6164 Seconds
(without alignments)
34.087 Million cell updates/sec

Title: A
Perfect score: 20
Sequence: 1 gseq 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 433172

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq.101002.*
2: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	22	AA60512
2	20	100.0	5	22	AA60535
3	20	100.0	6	22	AA60536
4	20	100.0	7	18	AAW34372
5	20	100.0	7	18	AAW10773
6	20	100.0	7	22	AA60507
7	20	100.0	8	14	AA632118
8	20	100.0	8	22	AA60537
9	20	100.0	9	22	AA60538
10	20	100.0	10	15	AA659475

11	20	100.0	10	22	AA695913
12	20	100.0	10	22	AA695943
13	20	100.0	10	22	AA695945
14	20	100.0	10	22	AA695981
15	20	100.0	10	22	AA685523
16	20	100.0	10	22	AA685525
17	20	100.0	10	22	AA686707
18	20	100.0	10	22	AA686709
19	20	100.0	10	22	AA686713
20	20	100.0	10	22	AA649023
21	20	100.0	11	22	AA620100
22	20	100.0	12	22	AA620389
23	20	100.0	13	22	AA670985
24	20	100.0	13	22	AA670986
25	20	100.0	13	22	AA670987
26	20	100.0	13	22	AB667621
27	20	100.0	14	20	AA633069
28	20	100.0	14	23	AB667529
29	20	100.0	15	17	AA649181
30	20	100.0	16	19	AA620460
31	20	100.0	17	18	AA608917
32	20	100.0	17	23	AA689904
33	20	100.0	17	23	AA689983
34	20	100.0	18	23	AA618035
35	20	100.0	19	22	AA662656
36	20	100.0	20	17	AA666926
37	20	100.0	20	22	AA660529
38	20	100.0	20	23	AA683897
39	20	100.0	21	22	AA604151
40	20	100.0	21	22	AA660525
41	20	100.0	21	23	AB664319
42	20	100.0	21	23	AA688211
43	20	100.0	21	23	AA698590
44	20	100.0	22	22	AB639372
45	20	100.0	22	22	AB624159

ALIGNMENTS

RESULT 1	
AA60512	
ID	AA60512 standard; peptide; 4 AA.
XX	
AC	AA60512;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Ghrelin-like growth hormone secretagogue (GHS) core region, SEQ ID NO:8.
XX	
KW	Growth hormone secretagogue; GHS; ghrelin; core region;
KM	calcium concentration elevation; infant growth disorder;
KX	growth hormone deficiency.
XX	
OS	Rattus norvegicus.
OS	Homo sapiens.
OS	Sus scrofa.
OS	Bos taurus.
OS	Gallus gallus.
OS	Anguilla japonica.
OS	Oncorhynchus mykiss.
XX	
PN	WO200107475-A1.
XX	
PD	01-FEB-2001.
XX	
PF	24-JUL-2000; 2000WO-0P04907.
XX	
PR	23-JUL-1999; 99JP-0210002.
PR	29-NOV-1999; 99JP-0338841.
PR	26-APR-2000; 2000JP-0126623.
XX	
PA	(KANG/) KANGAWA K.

XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX WPI; 2001-159704/16.
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 XX
 PS Claim 3; Page 184; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
 XX
 SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 ||||
 Db 1 GSSF 4

RESULT 2
 AAB60535
 ID AAB60535 standard; peptide; 5 AA.

XX
 AC AAB60535;

XX 24-APR-2001 (first entry)

XX Ghrelin-like growth hormone secretagogue (GHS) core region peptide #1.

XX Growth hormone secretagogue; GHS; ghrelin; core region;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.

XX Rattus norvegicus.
 OS Homo sapiens.
 OS Sus scrofa.
 OS Bos taurus.
 OS Gallus gallus.
 OS Anguilla japonica.
 OS Oncorhynchus mykiss.

XX WO200107475-A1.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-JP04907.

XX 23-JUL-1999; 99JP-0210002.

XX 29-NOV-1999; 99JP-0338841.

XX 26-APR-2000; 2000JP-0126623.

XX (KANG/) KANGAWA K.

XX

PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX WPI; 2001-159704/16.
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 XX
 PS Disclosure; Page 7; 210pp; Japanese.

XX The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.

XX Sequence 5 AA;

Query Match 100.0%; Score 20; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 ||||
 Db 1 GSSF 4

RESULT 3

AAB60536

ID AAB60536 standard; peptide; 6 AA.

XX
 AC AAB60536;

XX 24-APR-2001 (first entry)

XX Ghrelin-like growth hormone secretagogue (GHS) core region peptide #2.

XX Growth hormone secretagogue; GHS; ghrelin; core region;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.

XX Rattus norvegicus.
 OS Homo sapiens.
 OS Sus scrofa.
 OS Bos taurus.
 OS Gallus gallus.
 OS Anguilla japonica.
 OS Oncorhynchus mykiss.

XX WO200107475-A1.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-JP04907.

XX 23-JUL-1999; 99JP-0210002.

XX 29-NOV-1999; 99JP-0338841.

XX 26-APR-2000; 2000JP-0126623.

XX (KANG/) KANGAWA K.

XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;

XX WPI, 2001-159704/16.
 DR
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 XX
 XX PS Disclosure; Page 7; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
 CC
 SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
 ||||
 Db 1 GSSF 4

RESULT 4

AAW34372
 ID AAW34372 standard; peptide; 7 AA.

XX AC AAW34372;

DT 04-MAR-1998 (first entry)

XX PKB substrate #19.

XX Protein kinase B; PKB; substrate; glycogen metabolism; pancreatic cancer;

KW regulator; protein synthesis; enzyme modulator; type II diabetes;
 KW insulin-stimulated crosslinked kinase; breast cancer; ovarian cancer;
 KW therapy.

XX OS Synthetic.

XX PN WO9722360-A2.

XX PD 26-JUN-1997.

XX PF 20-DEC-1996; 96WO-GB03186.

XX PR 18-JUL-1996; 96GB-0015066.

XX PR 20-DEC-1995; 95GB-0026083.

XX PR 16-MAY-1996; 96GB-0010272.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI (UYDU-) UNIV DUNDEE.

XX PI Aleesi D, Cohen P, Cross D;

XX WPI, 1997-341435/31.

PT Use of protein kinase B for regulation of glycogen metabolism and
 PT protein synthesis - also peptide substrates for PKB and methods for

PT screening for modulators
 XX
 XX PS Example 8; Page 58; 98pp; English.
 XX

CC This sequence represents a substrate for protein kinase B (PKB). The use
 CC of PKB, its analogues, isoforms, inhibitors, activators and/or functional
 CC equivalents for regulating glycogen metabolism and/or protein synthesis
 CC is the subject of the invention. This sequence can also be used in a
 CC method of the invention for identifying agents that modulate the activity
 CC of PKB. It can also be used to screen for modulators of enzymes that
 CC catalyse PKB phosphorylation. PKB (an insulin-stimulated crosslinked
 CC kinase) and its analogues etc. are used to treat disease characterised by
 CC abnormal glycogen metabolism and/or protein synthesis, especially
 CC type II diabetes and cancer (specifically of breast, pancreas and ovary).
 CC The various screening methods are used to identify agents potentially
 CC useful for treating these diseases.
 CC
 SQ Sequence 7 AA;

Query Match 100.0%; Score 20; DB 18; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
 ||||
 Db 4 GSSF 7

RESULT 5

AAW10773
 ID AAW10773 standard; peptide; 7 AA.

XX AC AAW10773;

DT 10-NOV-1997 (first entry)

XX Ferritin motif #16 important for selective binding affinity.

KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
 KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
 KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
 KW pregnancy; infectious disease; ferritin; myosin light chain; tropoin;
 KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
 KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
 KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
 KW Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
 KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.

XX OS Synthetic.

XX PN WO9641172-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US10498.

XX PR 07-JUN-1995; 95US-0476375.

XX PA (CYTO-) CYTOGEN CORP.

XX PI Carter JM, Lee-Own FV;

XX WPI, 1997-077284/07.

PT Labelled functional surrogate of an analyte - useful as competitor
 PT molecule in affinity assays, esp. for detecting large macromolecules
 PT such as ferritin

XX PS Claim 56; Page 55; 156pp; English.

XX This sequence represents a peptide motif derived from ferritin which is
 CC important for selective binding affinity. Peptides containing motifs
 CC such as this may be used as functional surrogates in the conjugate of

the invention. The novel labelled conjugate comprises at least one label attached to a functional surrogate of an analyte of interest. The surrogate is capable of competing effectively with the analyte for a limiting amount of an affinity receptor for the analyte. The conjugate exhibits an activity that is altered upon interaction with the affinity receptor and this activity can be measured and related to the amount of the analyte present in a sample. Functional surrogates such as this have an immunoreactive group that allows the surrogate to compete effectively and with the analyte for a limiting amount of its affinity receptor. Functional surrogates are able to mimic naturally occurring analyses. They can be labelled for use in standard competitive affinity assays (esp. homogenous immunoassays) for detecting large macromolecules such as polypeptides, polysaccharides, polynucleotides, glycoproteins and lipid-containing macromolecules, as well as small haptens. Typical diagnostic analytes for detection include cardiac or tumour markers, allergens, hormones related to fertility-pregnancy or analyses associated with infectious disease. In particular, the assays are useful for detecting ferritin, follicle stimulating hormone, human growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human placental lactogen, hepatitis antigens or antibodies against them, human chorionic gonadotropin, human luteinising hormone, cytomegalovirus, Chlamydia, Streptococcus A, rubella, toxoplasma, herpesvirus, DK-MB, myoglobin, myosin light chain, troponin, carcinoembryonic antigen, alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).

Sequence 7 AA;

Query Match 100.0%; Score 20; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|
|
|
|
Db 3 GSSF 6

RESULT 6

AAB60507.
ID AAB60507 standard; peptide; 7 AA.
XX
AC AAB60507;
XX
DT 24-APR-2001 (first entry)
XX
DE Ghrelin-like growth hormone secretagogue (GHS) core region, SEQ ID NO:1.
XX
KW Growth hormone secretagogue; GHS; ghrelin; core region;
KW calcium concentration elevation; infant growth disorder;
KW growth hormone deficiency.
OS Rattus norvegicus.
OS Homo sapiens.
OS Sus scrofa.
OS Bos taurus.
OS Gallus gallus.
OS Anquilla japonica.
OS Oncomorhynchus mykiss.
XX
PN WO200107475-A1.
XX
PD 01-FEB-2001.
XX
PF 24-JUL-2000; 2000WO-JP04907.
XX
PR 23-JUL-1999; 99JP-0210002.
PR 29-NOV-1999; 99JP-0338841.
PR 26-APR-2000; 2000JP-0126623.
XX
PA (KANG/) KANGAWA K.
XX
PI Kangawa K, Kojima M, Hosoda H, Matsu H, Minamitake Y;
XX
XX WPI; 2001-159704/16.
DR

XX

PT New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
PS Disclosure; Page 180; 210pp; Japanese.

XX The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.

Sequence 7 AA;

Query Match 100.0%; Score 20; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|
|
|
|
Db 1 GSSF 4

RESULT 7

AAR32118
ID AAR32118 standard; Protein; 8 AA.
XX
AC AAR32118;
XX
DT 28-MAY-1993 (first entry)
XX
DE Heavy chain CDR3 region for humanised anti-IL-2 receptor antibody.
XX
KW antibody; monoclonal; chimaeric; grafted; humanised; IL-2;
KW interleukin-2; cytokines; interleukin-2 receptor; 55kD beta chain;
KW activated T cells; T cell mediated disease; graft versus host disease;
KW transplant rejection; autoimmune diseases; chemotherapy;
KW immunosuppressants; T cell typing; diagnosis; testing; detection; ss.
XX
OS Rattus rattus.
XX
PN WO9301289-A.
XX
PD 21-JAN-1993.
XX
PF 10-JUL-1992; 92WO-GB01258.
XX
PR 11-JUL-1991; 91GB-0015010.
XX
PA (WALD/) WALDMANN H.
PA (WELL) WELLCOME FOUND LTD.
XX
PI Crowe JS, Lewis AP, Waldmann H, Winter GP;
XX
XX WPI; 1993-045493/05.
DR N-PSDB; AAQ36579.
XX
XX Human interleukin-2 receptor antibodies - useful for treating and
XX preventing T-cell mediated diseases e.g. graft versus host
XX disease, transplant rejection etc.

PS Claim 1; Page 34; 48bp; English.
XX
CC This sequence represents the heavy chain CDR3 region from rat monoclonal
CC antibody YTH 906.9.25 which binds to the 55kD beta-chain of the IL-2
CC receptor on activated T cells. The encoding DNA was used in the
CC construction of humanised anti-IL2 receptor by PCR method using
CC CMPEPR-1H L chain as light chain template, and a humanised anti-CD4
CC heavy chain as a template. The humanised anti-IL2 receptor Ab can be
CC used in the treatment of T-cell mediated diseases eg. graft versus host
CC disease, transplant rejection, and various autoimmune diseases. It
CC may be administered alone or with chemotherapeutic or
CC immunosuppressive agents. In addn. it can be used for T cell
CC typing, to isolate specific IL-2R bearing cells and for diagnosis.
SQ Sequence 8 AA;
Query Match 100.0%; Score 20; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSSF 4
Db 1 GSSF 4
RESULT 8
AAB60537
ID AAB60537 standard; peptide; 8 AA. —
AC AAB60537;
XX 24-APR-2001 (first entry)
DT
XX Ghrelin-like growth hormone secretagogue (GHS) core region peptide #3.
DE
XX Growth hormone secretagogue; GHS; ghrelin; core region;
KW calcium concentration elevation; infant growth disorder;
KV growth hormone deficiency.
XX
OS Rattus norvegicus.
OS Homo sapiens.
OS Sus scrofa.
OS Bos taurus.
XX
PN WO200107475-A1.
PD 01-FEB-2001.
XX 24-JUL-2000; 2000WO-JP04907.
PF
XX 23-JUL-1999; 99JP-0210002.
PR 29-NOV-1999; 99JP-0338841.
PR 26-APR-2000; 2000JP-0126623.
XX
PA (KANG/) KANGAWA K.
XX
PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
XX WPI; 2001-159704/16.
DR
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
PS Disclosure; Page 7; 210pp; Japanese.
XX
CC The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC

CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting for a peptide. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
SQ Sequence 8 AA; —
Query Match 100.0%; Score 20; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSSF 4
Db 1 GSSF 4
RESULT 9
AAB60538
ID AAB60538 standard; peptide; 9 AA. —
AC AAB60538;
XX 24-APR-2001 (first entry)
DT
XX Ghrelin-like growth hormone secretagogue (GHS) core region peptide #4.
DE
XX Growth hormone secretagogue; GHS; ghrelin; core region;
KW calcium concentration elevation; infant growth disorder;
KV growth hormone deficiency.
XX
OS Rattus norvegicus.
OS Homo sapiens.
OS Sus scrofa.
OS Bos taurus.
XX
PN WO200107475-A1.
PD 01-FEB-2001.
XX 24-JUL-2000; 2000WO-JP04907.
PF
XX 23-JUL-1999; 99JP-0210002.
PR 29-NOV-1999; 99JP-0338841.
PR 26-APR-2000; 2000JP-0126623.
XX
PA (KANG/) KANGAWA K.
XX
PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
XX WPI; 2001-159704/16.
DR
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
PS Disclosure; Page 7; 210pp; Japanese.
XX
CC The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases

CC caused by a deficiency in growth hormone expression or activity. In particular, they are useful for promoting infant growth due to growth hormone deficiency. The compounds of the invention are safe with CC no accompanying side effects. The present sequence represents a CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.


XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|
|
|
|
Db 1 GSSF 4

RESULT 10

AAR59475

ID AAR59475 standard; peptide; 10 AA. 

AC AAR59475;

XX 26-JAN-1995 (first entry)

DE Neuropeptide for controlling muscle contraction.

XX Neuropeptide; muscle; muscular; contraction.

OS Helix pomatia.

XX JP06100590-A.

PN 12-APR-1994.

XX 21-SEP-1992; 92JP-0293615.

XX 21-SEP-1992; 92JP-0293615.

XX (SUNR) SUNTORY LTD.

XX WPI; 1994-156654/19.

PT Peptide having specific aminoacid sequence at its C-terminus - is
PI used to control muscular contraction

XX Claim 1; Page 2; 8pp; Japanese.

CC The peptide, which can be isolated from the ganglion of Helix
CC pomatia or can be synthesised, has the activity to improve muscular
CC contraction, as demonstrated by tests using anterior bysusus retractor
CC muscle.

XX Sequence 10 AA;

Query Match 100.0%; Score 20; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|
|
|
|
Db 6 GSSF 9

RESULT 11

AAG95913

ID AAG95913 standard; Peptide; 10 AA.

AC AAG95913;

XX 18-SEP-2001 (first entry)

DE Human complementary peptide, SEQ ID NO: 2107.

XX Human; complementary peptide; ligand; drug discovery; drug design.
XX Homo sapiens.
XX WO200142277-A2.
XX 14-JUN-2001.
XX 13-DEC-2000; 2000WO-GB04776.
XX 13-DEC-1999; 99GB-0029464.
XX (PROT-) PROTEOM LTD.
XX Roberts GW, Heal JR;
XX WPI; 2001-408419/43.
XX A set of peptide ligands consisting of specific complementary peptides
XX to proteins encoded by genes of the human genome, useful in an assay
XX for screening and identifying of one or more novel peptides which are
XX drug candidates or pro-drugs -
XX Example 4; Page 344; 646pp; English.
XX The invention relates to a set of complementary peptide ligands
XX generated from the human genome. The complementary peptides
XX interact with their relevant target proteins encoded in the human
XX genome. They can be used as reagents in drug discovery and as lead
XX ligands to facilitate drug design and development. The present
XX sequence is a complementary peptide provided in the specification.
XX Sequence 10 AA;
Query Match 100.0%; Score 20; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSSF 4
|
|
|
|
Db 5 GSSF 8
RESULT 12
AAG95943
ID AAG95943 standard; Peptide; 10 AA.
XX AAG95943;
XX 18-SEP-2001 (first entry)
XX Human complementary peptide, SEQ ID NO: 2137.
XX Human; complementary peptide; ligand; drug discovery; drug design.
XX Homo sapiens.
XX WO200142277-A2.
XX 14-JUN-2001.
XX 13-DEC-2000; 2000WO-GB04776.
XX 13-DEC-1999; 99GB-0029464.
XX (PROT-) PROTEOM LTD.
XX Roberts GW, Heal JR;
XX WPI; 2001-408419/43.
XX A set of peptide ligands consisting of specific complementary peptides

PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -

XX Example 4; Page 348; 646bp; English.

CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query March 100.0%; Score 20; DB 22; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4

DB 6 GSSF 9

RESULT 13

AAG5945 ID AAG5945 standard; Peptide; 10 AA.

AC AAG5945;

DT 18-SEP-2001 (first entry)

DE Human complementary peptide; SEQ ID NO: 2139.

KM Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB04776.

PR 13-DEC-1999; 99GB-0029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI, 2001-408419/43.

PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -

XX Example 4; Page 349; 646bp; English.

CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query March 100.0%; Score 20; DB 22; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4

DB 7 GSSF 10

RESULT 14

AAG5981 ID AAG5981 standard; Peptide; 10 AA.

AC AAG5981;

DT 18-SEP-2001 (first entry)

DE Human complementary peptide; SEQ ID NO: 2175.

KM Human; complementary peptide; ligand; drug discovery; drug design.

OS Homo sapiens.

PN WO200142277-A2.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB04776.

PR 13-DEC-1999; 99GB-0029464.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI, 2001-408419/43.

PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -

XX Example 4; Page 354; 646bp; English.

CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query March 100.0%; Score 20; DB 22; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4

DB 6 GSSF 9

RESULT 15

AAG5523 ID AAG5523 standard; Peptide; 10 AA.

AC AAG5523;

DT 11-SEP-2001 (first entry)

DE Saccharomyces cerevisiae peptide; SEQ ID NO: 472.

KM Saccharomyces cerevisiae; complementary peptide; peptide identification;

OS drug discovery; drug design.

XX Saccharomyces cerevisiae.

PN WO200142276-A1.

```

XX
PD 14-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-GB04773.
XX
XX 13-DEC-1999; 99GB-0029471.
XX
XX (PROT-) PROTEOM LTD.
XX
XX Roberts GW, Heal JR;
XX
XX WPI; 2001-367863/38.
XX
XX Identifying complementary peptides by analysis of protein and
XX nucleotide sequence databases, useful in drug design -
XX
XX Example 3; Page 95; 488pp; English.
XX
XX The invention relates to the identification of complementary peptides
XX by analysis of protein and nucleotide sequence databases from higher
XX eukaryotic genomes, excluding human and plants. The specific
XX complementary peptides interact with their relevant target proteins
XX encoded in the eukaryote genome. The peptides may be used as reagents
XX and drugs for drug discovery and as lead ligands for drug design and
XX development. The present sequence is a complementary peptide from
XX Saccharomyces cerevisiae.
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 20; DB 22; Length 10;
Best Local Similarity 100.0%; Pred.No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
   |||
Db 1 GSSF 4

Search completed: January 10, 2003, 15:59:13
Job time : 17.6364 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 10, 2003, 15:55:16 ; Search time 6.36364 Seconds
(without alignments)
60.427 Million cell updates/sec

Title: A
Perfect score: 20
Sequence: 1 gssf 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 11827

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	25	2	PH1733
2	20	100.0	33	2	PH1738
3	20	100.0	33	2	PH1742
4	20	100.0	33	2	PH1739
5	20	100.0	38	2	A56814
6	20	100.0	39	2	I40555
7	20	100.0	39	2	T12912
8	20	100.0	42	2	C82342
9	20	100.0	49	2	B97874
10	17	85.0	12	2	S25056
11	17	85.0	12	2	A49261
12	17	85.0	13	2	S32473
13	17	85.0	15	2	I67525
14	17	85.0	18	2	S66627
15	17	85.0	22	2	B37711
16	17	85.0	23	2	PH1694
17	17	85.0	24	2	PH1685
18	17	85.0	24	2	PH1698
19	17	85.0	24	2	PH1712
20	17	85.0	24	2	PH1713
21	17	85.0	25	2	PH1700
22	17	85.0	25	2	PH1701
23	17	85.0	25	2	A49038
24	17	85.0	26	2	PH1702
25	17	85.0	26	2	PH1703
26	17	85.0	26	2	PH1704
27	17	85.0	26	2	PH1718
28	17	85.0	27	2	PH1719
29	17	85.0	29	2	B61613

30	17	85.0	29	2	A83923	hypothetical prote
31	17	85.0	31	2	B49038	Ig lambda chain V
32	17	85.0	32	2	C26889	T-cell receptor be
33	17	85.0	34	2	PH1747	Ig heavy chain V r
34	17	85.0	34	2	H81223	hypothetical prote
35	17	85.0	35	2	E38601	Ig kappa chain V r
36	17	85.0	37	1	S26087	plasmodium-plaet
37	17	85.0	37	2	B36511	hypothetical prote
38	17	85.0	38	2	C49038	Ig lambda chain V
39	17	85.0	39	2	A82707	hypothetical prote
40	17	85.0	43	2	T07153	ethylene-responsiv
41	17	85.0	43	2	C30518	Ig heavy chain V-A
42	17	85.0	46	2	S11913	probable nitrogen
43	17	85.0	47	1	MOBP57	gene 0.5 protein -
44	17	85.0	47	2	JT0518	Ig heavy chain V-I
45	17	85.0	48	2	S02208	osteocalcin - emu

ALIGNMENTS

RESULT 1

PH1733
Ig heavy chain V region (clone GCC-13) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1733
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1733
A:Molecule type: mRNA
A:Residues: 1-25 <MCH>
A:Experimental source: B cell
A>Note: The authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
DB 20 GSSF 23

RESULT 2

PH1738
Ig heavy chain V region (clone NP-12-4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1738
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1738
A:Molecule type: mRNA
A:Residues: 1-33 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
DB 28 GSSF 31

RESULT 3

PH1742
Ig heavy chain V region (clone NP-12-8) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1742
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385

A:Accession: PH1742
A:Molecule type: mRNA
A:Residues: 1-33 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
DB 28 GSSF 31

RESULT 4

PH1739
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1739; PH1737
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385

A:Accession: PH1739
A:Molecule type: mRNA
A:Residues: 1-33 <MCH>
A:Experimental source: B cell, clone NP-12-3
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
DB 28 GSSF 31

RESULT 5

A56814
peptidylprolyl isomerase (EC 5.2.1.8) [similarity] - mouse (fragments)
N:Alternate names: cyclophilin homolog Sip24
C:Species: Mus musculus (house mouse)
C>Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 02-Sep-2000
C:Accession: A56814
R:Davis, T.R.; Tabatabai, L.; Bruns, K.; Hamilton, R.T.; Nilsson-Hamilton, M.
Biochim. Biophys. Acta 1095, 145-152, 1991

A:Title: Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize and secrete

A:Reference number: A56814; MUID:92031730; PMID:1932134

A:Accession: A56814
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-38 <DAV>
A:Experimental source: BALB/c 3T3 cells

A:Note: sequence modified after extraction from NCBI backbone
C:Superfamily: peptidylprolyl isomerase; cyclophilin homology
C:Keywords: cis-trans-isomerase

Query Match 100.0%; Score 20; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
DB 18 GSSF 21

RESULT 6

I40555
rap60B protein - Bacillus subtilis plasmid pTA1040

C:Species: Bacillus subtilis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40555
R:Meijer, W.J.; Venema, G.; Bron, S.
Nucleic Acids Res. 23, 612-619, 1995
A:Title: Characterization of single strand origins of cryptic rolling-circle plasmids
A:Reference number: I40549; MUID:95206941; PMID:7899081

A:Accession: I40555
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-39 <RES>
A:Cross-references: EMBL:U32378; NID:g1049109; PIDN:AAC44412.1; PID:g1049116
A:Experimental source: plasmid pTA1040
C:Genetics:
A:Gene: rap60B
A:Genome: plasmid

Query Match 100.0%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
DB 20 GSSF 23

RESULT 7

T12912
hypothetical protein yosA - Bacillus subtilis phage SPBc2

C:Species: Bacillus subtilis phage SPBc2
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 15-Oct-1999
C:Accession: T12912; C69925
R:Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mausel, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997

A:Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 prophage
A:Reference number: Z17583

A:Accession: T12912
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-39 <LAZ>

A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025626; PIDN:AAC13121.1
R:Kunat, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A:Erlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; Tognoni, J.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gran-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: C69925
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-39 <KUN>
A:Cross-references: GB:299114; GB:AL009126; NID:g2634230; PIDN:CAB13911.1; PID:e185491;
A:Experimental source: strain 168
C:Genetics:
A:Gene: y08A

Query Match 100.0%; Score 20; DB 2; Length 39; —
Best Local Similarity 100.0%; Pred. No. 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 17 GSSF 20

RESULT 8
C82342
hypothetical protein VC0279 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82342
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.D.;
charadon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nucleotide 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: C82342
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-42 <HET>
A:Cross-references: GB:AB004116; GB:AE003852; NID:g9654687; PIDN:AAF93454.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0279
A:Map position: 1

Query Match 100.0%; Score 20; DB 2; Length 42; —
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 38 GSSF 41

RESULT 9
B97874
degenerate transposase (orf2) [imported] - *Streptococcus pneumoniae* (strain R6)
C:Species: *Streptococcus pneumoniae*
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: B97874
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Buggett, S.; Dehoff, B.S.; E.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; Mcahren, S.; M.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Author: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B97874
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-49 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK98822.1; PID:g15457547; GSPDB:GN00174
C:Genetics:
A:Gene: IS1167-truncation

Query Match 100.0%; Score 20; DB 2; Length 49; —
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 40 GSSF 43

RESULT 10
S25056
Ig heavy chain - mouse
C:Species: *Mus musculus* (house mouse)
C>Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 20-Jun-2000
C:Accession: S25056
R:Jacob, J.; Kelsoe, G.
submitted to the EMBL Data Library, July 1992
A:Description: In situ studies on the primary immune response to (4-hydroxy-3-nitrophenyl)
A:Reference number: S25024
A:Accession: S25056
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-12 <JAC>
A:Cross-references: EMBL:X67386; NID:g50927; PIDN:CAA47798.1; PID:g1333920
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 85.0%; Score 17; DB 2; Length 12; —
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 8 GSSF 11

RESULT 11
A49261
coagulation factor X inhibitor - sharp-nosed viper (fragment)
C:Species: *Agkistrodon acutus* (sharp-nosed viper)
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-May-1994
C:Accession: A49261
R:Cox, A.C.
Toxicol. 31, 1445-1457, 1993
A:Title: Coagulation factor X inhibitor from hundred-pace snake (*Deinagkistrodon acutus*)
A:Reference number: A49261; MUID:94143901; PMID:8310445
A:Accession: A49261
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <COX>

Query Match 85.0%; Score 17; DB 2; Length 12; —
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 6 GSSF 9

RESULT 12
S32473
Lymnaeidae 3 - great pond snail
C:Species: *Lymnaea stagnalis* (great pond snail)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S32473
R:Johnsen, A.H.; Rehfeld, J.F.
Eur. J. Biochem. 213, 875-879, 1993
A:Title: Lymnaeidae, a new family of neuropeptides from the pond snail, *Lymnaea stagna*
A:Reference number: S32473; MUID:93038777; PMID:8477756
A:Accession: S32473
A:Molecule type: protein
A:Residues: 1-13 <JOH>
A:Cross-references: PIDN:AA826364.1; PID:g299831
A:Experimental source: ganglia
C:Keywords: amidated carboxyl end; neuropeptide

F;13/Modified site: amidated carboxyl end (Phe) #status predicted

Query Match 85.0%; Score 17; DB 2; Length 13; ✓
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
Db 7 GSAP 10

RESULT 13

I67525

CD33 antigen homolog - mouse (fragment)

C;Species: Mus sp. (mouse)

C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Jun-1998

C;Accession: I67525

R;Chies, J.A.; Lembezat, M.P.; Freitas, A.A.

Eur. J. Immunol. 24, 1657-1664, 1994

A;Title: Entry of B lymphocytes into the persistent cell pool in non-immunized mice is r

A;Reference number: I53392; MUID:94298870; PMID:8026526

A;Accession: I67525

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-15 <RES>

A;Cross-references: GB:S71349; NID:9550037

C;Genetics:

A;Gene: Ig VH7183

Query Match 85.0%; Score 17; DB 2; Length 15;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
Db 9 GSSV 12

RESULT 14

S66627

tau protein - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997

C;Accession: S66627

R;Moreno, F.J.; Medina, M.; Perez, M.; Montejo de Garcini, E.; Avila, J.

FEBS Lett. 372, 65-68, 1995

A;Title: Glycogen synthase kinase 3 phosphorylates recombinant human tau protein at seri

A;Reference number: S66627; MUID:96032547; PMID:7556645

A;Accession: S66627

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-8;9-18 <MOR>

Query Match 85.0%; Score 17; DB 2; Length 18;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
Db 2 GSTP 5

RESULT 15

B32711

leghemoglobin - Lotus corniculatus (fragment)

C;Species: Lotus corniculatus

C;Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 04-Mar-2000

C;Accession: B32711

R;Stougaard, J.; Petersen, T.E.; Marcker, K.A.

Proc. Natl. Acad. Sci. U.S.A. 84, 5754-5757, 1987

A;Title: Expression of a complete soybean leghemoglobin gene in root nodules of transfer

A;Reference number: A32711

A;Accession: B32711

A;Molecule type: protein
A;Residues: 1-22 <STO>
C;Superfamily: globin; globin homology
C;Keywords: oxygen carrier

Query Match 85.0%; Score 17; DB 2; Length 22;
Best Local Similarity 75.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
|||
Db 11 GSSV 14

Search completed: January 10, 2003, 15:56:27

Job time : 7.36364 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:16 ; Search time 3.63636 Seconds
(without alignments)
45.624 Million cell updates/sec

Title: A
Perfect score: 20
Sequence: 1 gseq 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 3754

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	85.0	13	NP3_LYMST	P80180 lymnaea ste
2	17	85.0	29	CERB_CERCA	P36191 ceratitidis c
3	17	85.0	37	PETG_EUGGR	P30396 euglena gra
4	17	85.0	47	V05_BPT7	P03777 bacterioph
5	17	85.0	48	OSTC_DRONO	P15504 dromaeus no
6	17	85.0	49	R22A_MOUSE	P35285 mus musculi
7	16	80.0	8	LCK5_LEUMA	P19287 leucophaea
8	16	80.0	24	V01_BPAL3	P08766 bacterioph
9	16	80.0	36	C3L1_BOVIN	P30922 bos taurus
10	16	80.0	47	RK2X_CYACA	P41551 cyanidium c
11	16	80.0	48	YMCF_EMENT	P03885 emericella
12	15	75.0	11	CA42_LITCI	P82092 litorea cit
13	14	70.0	8	LCK6_LEUMA	P19288 leucophaea
14	14	70.0	12	CXL3_CONMR	P58809 conus maro
15	14	70.0	12	PVK2_PERAM	P81555 periplaneta
16	14	70.0	13	CPI_APLCA	Q10998 aplysia cal
17	14	70.0	15	FIBA_ANAPL	P12801 anae platyr
18	14	70.0	15	TAI_TREBR	P34070 tremella dr
19	14	70.0	15	TRPA_LEUMA	P81753 leucophaea
20	14	70.0	17	FLAW_AROCH	P23002 azotobacter
21	14	70.0	18	AHD2_TETPY	P35303 tetrahymena
22	14	70.0	20	FLAW_AZOV1	P52964 azotobacter
23	14	70.0	20	PSAK_PSA	P17226 pisum sativ
24	14	70.0	20	SB18_MAIZE	P82867 zea mays (m
25	14	70.0	21	LPRM_CORDI	P21332 corynebacte
26	14	70.0	22	13KD_BACST	P80166 bacillus st
27	14	70.0	24	FEDG_AMEYE	P80707 amycolatops
28	14	70.0	25	H11_WHEAT	P15871 triticum ae
29	14	70.0	25	PLRT_PSEAM	P81941 pseudopleur
30	14	70.0	26	CX06_CONTU	P58915 conus tulip
31	14	70.0	29	PSAK_SPTOL	P14627 splinacia ol
32	14	70.0	30	KAB5_OLDAP	P58456 oldenlandia
33	14	70.0	31	DEF2_MESAV	P81466 mesocricetu

34	14	70.0	32	1	UKA6_HUMAN	P31942 homo sapien
35	14	70.0	32	1	Y169_TREPA	O83199 treponema p
36	14	70.0	33	1	ALOX_PICPA	P04842 pichia past
37	14	70.0	33	1	DEF4_MESAV	P81466 mesocricetu
38	14	70.0	33	1	LPRH_ECOLI	P37324 escherichia
39	14	70.0	33	1	SC63_CANFA	P82008 canis fami
40	14	70.0	34	1	LPTN_PROVU	P28779 proteus vul
41	14	70.0	35	1	COPA_CANFA	P40765 canis fami
42	14	70.0	35	1	PETG_CYACA	Q9C199 cyanidium c
43	14	70.0	36	1	HIL5_ENSMI	P27203 ensis minor
44	14	70.0	36	1	NPF_ARTTR	P41334 artiposthi
45	14	70.0	37	1	F13A_BOVIN	P12260 bos taurus

ALIGNMENTS

```

RESULT 1
NP3_LYMST          STANDARD;          PRT;          13 AA.
AC  P80180;
DT  01-JUL-1993 (Rel. 26, Created)
DT  01-JUL-1993 (Rel. 26, Last sequence update)
DT  01-JUL-1993 (Rel. 26, Last annotation update)
DE  Lymnaea-Df-amide 3.
OS  Lymnaea stagnalis (Great pond snail).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC  Lymnaeidae; Lymnaea.
OX  NCBI_TaxId=6523;
RN  [1]
RP  LYMST
RC  TISSUE=ganglion;
RX  MEDLINE=93238777; PubMed=8477756;
RA  Johnsen A.H., Rehfeld J.F.;
RT  "Lymnaeidae, a new family of neuropterids from the pond snail,
RT  Lymnaea stagnalis. Clue to cholecystokinin immunoreactivity in
RT  invertebrates?";
RL  Eur. J. Biochem. 213:875-879(1993).
CC  -!- SIMILARITY: RELATED TO THE CHOLECYSTOKININ (CKK) FAMILY.
DR  PIR, S32473; S32473.
KW  Neuropeptide; Amidation.
FT  MOD_RES 13 13  AMIDATION.
FT  UNRES 12 12
SQ  SEQUENCE 13 AA; 1462 MW; 9CA07BA3F5D5B865 CRC64;

Query Match          Score 17; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 88;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 GSSF 4
Db  7 GSAF 10

RESULT 2
CERB_CERCA          STANDARD;          PRT;          29 AA.
AC  P36191;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-FEB-1996 (Rel. 33, Last annotation update)
DE  Ceratocoxin B.
CN  CTRB
OS  Ceratitidis capitata (Mediterranean fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC  Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC  Muscomorpha; Tephritidae; Tephritidae; Ceratitidis.
OX  NCBI_TaxId=7213;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=female accessory gland;
RX  MEDLINE=93557786; PubMed=8353519;
RA  Marchini D., Giordano P.C., Amans R., Bernini L.F., Dallai R.;

```

"Purification and primary structure of ceratotoxin A and B, two antibacterial peptides from the female reproductive accessory glands of the medfly *Ceratitis capitata* (Insecta:Diptera).";
 RL Insect Biochem. Mol. Biol. 23:591-598(1993).
 CC -!- FUNCTION: FEMALE-SPECIFIC PEPTIDES WITH POTENT ACTIVITY AGAINST GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA. THEY HAVE AS WELL HEMOLYTIC ACTIVITY. THESE PROTEINS ARE STABLE EVEN AT 100 DEGREES CELSIUS.
 CC -!- SUBUNIT: HOMOPOLYMER OF FOUR TO SIX SUBUNITS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: STRUCTURALLY RELATED TO CECROPINS, DEFENSINS AND APIADECINS.
 KW Insect immunity; Hemolysis; Antibiotic.
 SQ SEQUENCE 29 AA; 2861 MW; BE57F4EECB2DA6B0 CRC64;
 Query Match 85.0%; Score 17; DB 1; Length 29;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSF 4
 DB 3 GSAP 6
 RESULT 3
 ID PETG_EUGGR STANDARD; PRT; 37 AA.
 AC P30356;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome B6-F complex subunit V (Cytochrome b6f complex subunit petg).
 DE petg.
 GN PETG.
 OS Euglena gracilis.
 OG Chloroplast.
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI_TaxID=3039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z;
 RX MEDLINE=93347989; PubMed=8346031;
 RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A., Oreat B., Spielmann A., Stutz E.;
 RT "Complete sequence of Euglena gracilis chloroplast DNA.";
 RL Nucleic Acids Res. 21:3537-3544(1993).
 CC -!- FUNCTION: THE CYTOCHROME B6-F COMPLEX FUNCTIONS IN THE LINEAR CROSS-MEMBRANE TRANSPORT OF ELECTRONS BETWEEN PHOTOSYSTEM II AND I, AS WELL AS IN CYCLIC ELECTRON FLOW AROUND PHOTOSYSTEM I. PETG IS REQUIRED FOR EITHER THE STABILITY OR ASSEMBLY OF THE CYTOCHROME B6-F COMPLEX.
 CC -!- SUBCELLULAR LOCATION: Thylakoid membrane-associated.
 CC -!- SIMILARITY: BELONGS TO THE PETG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z11874; CAA77909.1; -;
 CC EMBL; X70810; CAA50092.1; -;
 CC PIR; S26087; S26087;
 CC PIR; S34513; S34513;
 CC InterPro; IPR003683; Cytochrmb6/f_5.
 CC Pfam; PF02529; Petg; 1.
 KW Electron transport; Chloroplast; Respiratory chain; Thylakoid; Transmembrane.
 FT DOMAIN 1 4 LUMENAL (POTENTIAL).
 FT TRANSMEM 5 25 POTENTIAL.
 FT DOMAIN 26 37 STROMAL (POTENTIAL).
 FT

SQ SEQUENCE 37 AA; 4147 MW; 13806339E110D3D6 CRC64;
 Query Match 85.0%; Score 17; DB 1; Length 37;
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSF 4
 DB 33 GNSF 36
 RESULT 4
 ID V05_BPT7 STANDARD; PRT; 47 AA.
 AC P03777;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-MAR-1989 (Rel. 10, Last annotation update)
 DE Gene 0.5 protein.
 DE 0.5.
 GN Bacteriophage T7.
 OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like viruses.
 OX NCBI_TaxID=10760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83241725; PubMed=6864790;
 RA Dunn J.J., Studier F.W.;
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements.";
 RL J. Mol. Biol. 166:477-535(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82078034; PubMed=7310871;
 RA Dunn J.J., Studier F.W.;
 RT "Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the beginning of gene 4";
 RL J. Mol. Biol. 148:303-330(1981).
 CC -!- FUNCTION: THE FUNCTION OF THIS EARLY GENE PROTEIN IS UNKNOWN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V01146; CAA24386.1; -;
 CC EMBL; V01127; CAA24329.1; -;
 CC PIR; A04402; W0BP57.
 CC PIR; S42285; S42285.
 CC SEQUENCE 47 AA; 4745 MW; B07BC5B9FC12FA66 CRC64;
 Query Match 85.0%; Score 17; DB 1; Length 47;
 Best Local Similarity 75.0%; Pred. No. 3.4e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSF 4
 DB 17 GASF 20
 RESULT 5
 ID OSTC_DRONO STANDARD; PRT; 48 AA.
 AC P15504;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).
 DE BGLAP.
 GN

```

OS Dromaeus novae-hollandiae (Emu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Casuariiformes; Dromiidae;
OC Dromaeus.
OX NCBI_Taxid=8790;
RN [1]
RP SEQUENCE.
RX MEDLINE=88134266; PubMed=3501719;
RA Hug N.L., Tseng A., Chapman G.E.;
RT "The amino acid sequence of Emu osteocalcin: gas phase sequencing of
  Gla-containing proteins.";
RL Biochem. Int. 15:271-277(1987).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTHrP-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC PIR: S02208; S02208.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS; PR00002; GLABONE.
DR SMART; SM00069; GLA; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxylglutamic acid; Vitamin K; Bone.
FT MOD_RES 16 16
FT MOD_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 23 23 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 22 28 BY SIMILARITY.
SQ SEQUENCE 48 AA; 5292 MW; 50A00F3BFA8C7FFD CRC64;

Query Match 85.0%; Score 17; DB 1; Length 48;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
DB 5 GSSF 8

RESULT 6
R22A_MOUSE STANDARD; PRT; 49 AA.
AC P35285;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RAB-related protein Rab-22A (RAB-14) (Fragment).
GN RAB22A OR RAB22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92210010; PubMed=1555775;
RA Chavrier P., Simons K., Zerial M.;
RT "The complexity of the Rab and Rho GTP-binding protein subfamilies
  revealed by a PCR cloning approach.";
RL Gene 112:261-264(1992).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; M79304; AAK14828.1; -.
DR PIR; JH0644; JH0644.
DR MGd; MG1:105072; Rab22.
DR InterPro: IPR001806; Ras_trnsfrmg.
DR Pfam; PF00071; ras; 1.
KW GTP-binding.
FT MOD_RES 1 1
FT NP_BIND <1 3 GTP (BY SIMILARITY).
FT NP_BIND 44 48 GTP (BY SIMILARITY).
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5666 MW; 7356DD677BB60F057 CRC64;

Query Match 85.0%; Score 17; DB 1; Length 49;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
DB 23 GSSF 26

RESULT 7
LCK5_LEUMA STANDARD; PRT; 8 AA.
AC P19987;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Leucokinin V (L-V).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberidae; Blaberidae; Leucophaea.
OX NCBI_Taxid=6988;
RN [1]
RP SEQUENCE.
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
  myotropic peptides of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:27-30(1987).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
DR PIR; J50315; J50315.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 80.0%; Score 16; DB 1; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GSSF 4
DB 1 GSSF 4

RESULT 8
VGJ_BPAL3 STANDARD; PRT; 24 AA.
AC P08766;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small core protein (J protein).
GN J.
OS Bacteriophage alpha-3, and
OC Bacteriophage phi-K.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
OX NCBI_Taxid=10849; 10848;
RN [1]

```

SEQUENCE FROM N.A.
RC SPECIES=Phage alpha-3;
RX MEDLINE=92223109; PubMed=1532908;
RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
RT "Nucleotide sequence of the genome of the bacteriophage alpha 3;
interrelationship of the genome structure and the gene products with
those of the phages, phi X174, G4 and phi K.";
RL Biochim. Biophys. Acta 1130:277-288(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Phage alpha-3;
RX MEDLINE=84294906; PubMed=6088949;
RA Kodaira K.-I., Taketo A.;
RT "Isolation and some properties of bacteriophage alpha3 gene J
mutant.";
RL Mol. Gen. Genet. 195:541-543(1984).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Phage phi-K;
RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases;
CC -!- FUNCTION: THE J PROTEIN IS ASSOCIATED WITH THE DNA AND IS SITUATED
IN AN INTERIOR CLEFT OF THE F PROTEIN.
CC -!- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
J PROTEINS, AND 12 COPIES OF THE H PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60322; CAA42880.1; -;
DR EMBL; X00774; CAA25349.1; -;
DR EMBL; X60323; CAA42890.1; -;
DR PIR; S09546; S09546.
DR PIR; S22333; S22333.
KW Coat protein; DNA-binding.
SQ SEQUENCE 24 AA; 2823 MW; 0EE261CFF11F669B CRC64;
Query Match 80.0%; Score 16; DB 1; Length 24;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSF 4
DB 21 GSQP 24
RESULT 9
C3L1 BOVIN ID - C3L1 BOVIN STANDARD; PRT; 36 AA.
AC P30922;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 (Cartilage glycoprotein-39) (GP-39) (39 kDa
whey protein) (Fragment).
GN CHI3L1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE
RX MEDLINE=88106603; PubMed=3122754;
RA Rejman J.J., Hurley W.L.;
RT "Isolation and characterization of a novel 39 kilodalton whey protein
from bovine mammary secretions collected during the nonlactating
period.";

RL Biochem. Biophys. Res. Commun. 150:329-334(1988).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY SECRETIONS COLLECTED DURING THE
NONLACTATING PERIOD.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
DR PIR; A27682; A27682.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1_18.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; PARTIAL.
KW Glycoprotein.
FT NON_TER 36 36
SQ SEQUENCE 36 AA; 4264 MW; 0FF5730DFF2E14A9 CRC64;
Query Match 80.0%; Score 16; DB 1; Length 36;
Best Local Similarity 75.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSF 4
DB 18 GSXF 21
RESULT 10
RK2X CYACA ID - RK2X CYACA STANDARD; PRT; 47 AA.
AC P4151;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chloroplast 50S ribosomal protein L27 (Fragment).
GN RPL27.
OS Cyanidium caldarium.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RK-1;
RX MEDLINE=94154241; PubMed=8111025;
RA Fujiwara S., Kawachi M., Inouye I., Someya J.;
RT "The gene for ribosomal protein L27 is located on the plastid rather
than the nuclear genome of the chlorophyll c-containing alga
Pleurochrysis carterae.";
RT Plant Mol. Biol. 24:253-257(1994).
CC -!- SIMILARITY: BELONGS TO THE L27P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26098; BAA05094.1; -;
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR ProDom; PD003114; Ribosomal_L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Chloroplast.
FT NON_TER 1 1
FT NON_TER 47 47
SQ SEQUENCE 47 AA; 5081 MW; C93F703848741964 CRC64;
Query Match 80.0%; Score 16; DB 1; Length 47;
Best Local Similarity 75.0%; Pred. No. 6.4e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSGF 4
DB 34 GSGF 37

RESULT 11
YVCF_EMBENI STANDARD; PRT; 48 AA.
ID_YVCF_EMBENI
AC P03885;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 5.6 kDa protein in COX1 intron (URE-F).
OS *Emicella nidulans* (*Aspergillus nidulans*).
OC Mitochondrion.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Euceliales; Trichocomaceae; *Emicella*.
OX NCBI_TaxId=5072;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=81135863; PubMed=7008953;
RA Koechel H.G., Lazarus C.M., Basak N., Kuentzel H.;
RT "Mitochondrial tRNA gene clusters in *Aspergillus nidulans*:
organization and nucleotide sequence."/;
RL Cell 23:625-633 (1981).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=83038633; PubMed=6290989;
RA Netzer R., Koechel H.G., Basak N., Kuentzel H.;
RT "Nucleotide sequence of *Aspergillus nidulans* mitochondrial genes
coding for ATPase subunit 6, cytochrome oxidase subunit 3, seven
unidentified proteins, four tRNAs and L-rRNA."/;
RL Nucleic Acids Res. 10:4783-4794 (1982).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: J01390; AAA99211.1;
DR EMBL: X07795; CAA30642.1;
DR PIR: A04518; OXASF.
KW Hypothetical protein; Mitochondrion.
SQ SEQUENCE 48 AA; 5628 MW; FCGABBD1CD1992DD CRC64;

Query Match Score 80.0%; Score 16; DB 1; Length 48;
Best Local Similarity 75.0%; Pred. No. 6.5e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSGF 4
DB 5 GSGF 8

RESULT 12
CA42_LITCI STANDARD; PRT; 11 AA.
ID_CA42_LITCI
AC P82052;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caerulain 4.2/4.2Y4.

OS *Litoria citropa* (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eueleostomi;
OC Amphibia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxId=94770;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.

RC TISSUE=Skin secretion;
RX MEDLINE=20057701; PubMed=10589099;
RA Wabnitz P.A., Bowie J.H., Tyler M.J.;
RT "Caerulain-like peptides from the skin glands of the Australian blue
RT montains tree frog *Litoria citropa*. Part 1. Sequence determination
RT using electrospray mass spectrometry."/;
RL Rapid Commun. Mass Spectrom. 13:2498-2502 (1999).
CC -1- FUNCTION: HYPOTENSIVE NEUROPEPTIDE (PROBABLE).
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- PTM: ISOFORM 4.2Y4 DIFFERS FROM ISOFORM 4.2 IN NOT BEING
CC SUPDATED.
CC -1- MASS SPECTROMETRY: MW=1404; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR Interpro: IPR001651; Gastrin.
DR PROSITE: PS00259; GASTRIN; FALSE NEG.

KW Amphibian skin; Hypotensive agent; Amidation; Sulfation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 4 4 SULFATION.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1344 MW; 10DAB894F5B861BB CRC64;

Query Match Score 75.0%; Score 15; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSGF 4
DB 6 GSGF 9

RESULT 13
LCK6_LEUMA STANDARD; PRT; 8 AA.
ID_LCK6_LEUMA
AC P19988;

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Leucokinin VI (L-VI).

OS *Leucophaea maderae* (*Madeira cockroach*).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberidae; Blaberidae; *Leucophaea*.
OX NCBI_TaxId=6988;
RN [1]

RP SEQUENCE.
RC TISSUE=Head;
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of *Leucophaea maderae*."/;
RL Comp. Biochem. Physiol. 88C:27-30 (1987).

CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTEIN (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS, AND TO MANDUCA SEXTA AND
CC HELIOTHIS ZEA ADIPOKINETIC HORMONE.

DR PIR: JS0316; JS0316
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 935 MW; 9D6365B1E9D5A5A6 CRC64;

Query Match Score 70.0%; Score 14; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSF 4
DB 2 SSF 4

RESULT 14
CXL3_CONMR STANDARD; PRT; 12 AA.
ID_CXL3_CONMR

AC PS8809;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin Mrx.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=1262.77; MW ERR=0.07; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT DISULFID 4 11
FT MOD_RES 11 11 HYDROXYLATION.
SQ SEQUENCE 12 AA; 1251 MW; 277AAE2422D5A2C8 CRC64;

Query Match 70.0%; Score 14; DB 1; Length 12;
Best Local Similarity 75.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 5 GVSF 8

RESULT 15
PVK2 PERAM
ID PVK2 PERAM STANDARD; PRT; 12 AA.
AC P8155;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periviscerokinin-2 (Pea-PVK-2).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattellidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Abdominal perisymphathetic organs;
RX MEDLINE=98326577; PubMed=9663444;
RA Predel R., Rapus J., Eckert M., Holman G.M., Nachman R.J., Wang Y.,
RA Penzlin H.;
RT "Isolation of periviscerokinin-2 from the abdominal perisymphathetic
RT organs of the American cockroach, Periplaneta americana.";
RL Peptides 19:801-809(1998).
CC -1- FUNCTION: MYOACTIVE PEPTIDE; HAS EXCITATORY ACTIONS ON THE
CC -1- HYPERNEURAL MUSCLE.
CC -1- MASS SPECTROMETRY: MW=1189.3; METHOD=MALDI.
KW Neuropeptide; Amidation.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1190 MW; 2F4D8EE1EB05728 CRC64;

Query Match 70.0%; Score 14; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSS 3

Db 1 GSS 3
Search completed: January 10, 2003, 15:55:46
Job time : 5.63636 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 10, 2003, 15:55:17 ; Search time 11.8182 Seconds
(without alignments)
69.739 Million cell updates/sec

Title: A
Perfect score: 20
Sequence: 1 gseq 4

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 33835

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_protist:*
11: sp_rotent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	17	11 Q9QVH8	Q9QVH8 mus sp. sup
2	20	100.0	34	6 Q8SPN7	Q8SPN7 macaca mula
3	20	100.0	38	10 Q9FJH1	Q9FJH1 arabidopsis
4	20	100.0	39	2 Q45446	Q45446 bacillus su
5	20	100.0	39	9 Q64160	Q64160 bacterioph
6	20	100.0	39	11 Q9R137	Q9R137 mus musculu
7	20	100.0	39	16 Q1888	Q1888 bacillus su
8	20	100.0	42	16 Q9K109	Q9K109 uncultured
9	20	100.0	47	3 Q8X109	Q8X109 vibrio chol
10	17	85.0	20	2 Q9R4G2	Q9R4G2 vibrio. bet
11	17	85.0	20	3 P82288	P82288 acromonium
12	17	85.0	20	6 Q9TQY9	Q9TQY9 sus scrofa
13	17	85.0	21	11 Q9R204	Q9R204 mus musculu
14	17	85.0	24	12 Q9W1E1	Q9W1E1 citrus tris
15	17	85.0	27	15 Q9E8R0	Q9E8R0 human immun
16	17	85.0	27	15 Q9E8Q9	Q9E8Q9 human immun

17	17	85.0	27	15	Q900F1	Q900F1 human immun
18	17	85.0	27	15	Q900F0	Q900F0 human immun
19	17	85.0	28	12	Q9YN93	Q9YN93 citrus tris
20	17	85.0	29	16	Q9KAV1	Q9KAV1 bacillus ha
21	17	85.0	30	2	Q68187	Q68187 lactococcus
22	17	85.0	34	15	Q79984	Q79984 human immun
23	17	85.0	34	15	Q79983	Q79983 human immun
24	17	85.0	34	16	Q9K1D1	Q9K1D1 neisseria m
25	17	85.0	35	15	Q90520	Q90520 human immun
26	17	85.0	35	15	Q9J4E0	Q9J4E0 human immun
27	17	85.0	35	15	Q79997	Q79997 human immun
28	17	85.0	35	15	Q79996	Q79996 human immun
29	17	85.0	35	15	Q77923	Q77923 human immun
30	17	85.0	35	15	Q77925	Q77925 human immun
31	17	85.0	35	15	Q77926	Q77926 human immun
32	17	85.0	35	15	Q77924	Q77924 human immun
33	17	85.0	35	15	Q77927	Q77927 human immun
34	17	85.0	35	15	Q77928	Q77928 human immun
35	17	85.0	35	15	Q77913	Q77913 human immun
36	17	85.0	35	15	Q77918	Q77918 human immun
37	17	85.0	35	15	Q77915	Q77915 human immun
38	17	85.0	35	15	Q77920	Q77920 human immun
39	17	85.0	35	15	Q77922	Q77922 human immun
40	17	85.0	35	15	Q77901	Q77901 human immun
41	17	85.0	35	15	Q77903	Q77903 human immun
42	17	85.0	35	15	Q77890	Q77890 human immun
43	17	85.0	35	15	Q77892	Q77892 human immun
44	17	85.0	35	15	Q77893	Q77893 human immun
45	17	85.0	35	15	Q77894	Q77894 human immun

ALIGNMENTS

RESULT 1
Q9QVH8 PRELIMINARY; PRT; 17 AA.
AC Q9QVH8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE SUPERINDUCIBLE protein 24, SIP24=CYCLOPHILIN homolog, PEAK A (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP MEDLINE=92031730; PubMed=1932134;
RX Davis T.R., Tabatabai L., Bruns K., Hamilton R.T., Nilsen-Hamilton M.; "Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize RT and secrete a cyclophilin-like protein and beta 2-microglobulin.";
RL Biochim. Biophys. Acta 1095:145-152(1991).
DR HSSP: P05092; 2CPL.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; pro_isomerase; 1.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1785 MW; 11276657FEB240D9 CRC64;
Query Match 100.0%; Score 20; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GSSF 4
Db 13 GSSF 16
RESULT 2
Q8SPN7 PRELIMINARY; PRT; 34 AA.
ID Q8SPN7

```

AC Q8SPN7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Slit-like protein 2 (Fragment).
GN SLIT2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Norgren R.B. Jr., Zink M.A., Jia Y., Ojeda S.R., Spindel E.R.;
RI "Construction of a targeted rhesus macaque microarray.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY083584; AAM11999.1; -.
FT NON_TER 1 34
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 3916 MW; E043A2D43BEE9134 CRC64;

Query Match 100.0%; Score 20; DB 6; Length 34; —
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 15 GSSF 18

RESULT 3
Q9FJBI
ID Q9FJBI PRELIMINARY; PRT; 38 AA.
AC Q9FJBI;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Genomic DNA, chromosome 5, Pl clone: MOK9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asanizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:297-308(1998).
DR EMBL; AB015477; BAB08707.1; -.
SQ SEQUENCE 38 AA; 4298 MW; 0340720C8BFB476B CRC64;

Query Match 100.0%; Score 20; DB 10; Length 38; —
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 6 GSSF 9

RESULT 4
Q45446
ID Q45446 PRELIMINARY; PRT; 39 AA.
AC Q45446;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE RAP60B.

```

```

GN RAP60B.
OS Bacillus subtilis.
OG Plasmid pTA1040.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM1232;
RX MEDLINE=95206941; PubMed=7899081;
RA Meijer W.J., Venema G., Bron S.;
RT "Characterization of single strand origins of cryptic rolling-circle
RT plasmids from Bacillus subtilis.";
RL Nucleic Acids Res. 23:612-619(1995).
DR EMBL; U32378; AAC44412.1; -.
KW Plasmid.
SQ SEQUENCE 39 AA; 4125 MW; C086DA26EB01AE03 CRC64;

Query Match 100.0%; Score 20; DB 2; Length 39; —
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 20 GSSF 23

RESULT 5
O64160
ID O64160 PRELIMINARY; PRT; 39 AA.
AC O64160;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 4.1 kDa protein.
GN YOSA.
OS Bacteriophage SPBc2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=66797;
RN [1]
RP SEQUENCE FROM N.A.
RA Lazarevic V., Duesterhoeft A., Soldo B., Hilbert H., Maue C.,
RA Karamata D.;
RT "The complete nucleotide sequence of the Bacillus subtilis SPbetac2
RT prophage.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020713; AAC13121.1; -.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4090 MW; AC22034F15E873E1 CRC64;

Query Match 100.0%; Score 20; DB 9; Length 39; —
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSF 4
Db 17 GSSF 20

RESULT 6
Q9R137
ID Q9R137 PRELIMINARY; PRT; 39 AA.
AC Q9R137;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cyclophilin A (Fragment).
GN PP1A OR CYPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=20422670; PubMed=10964515;
RA Colgan J., Asmal M., Luban J.;
RT "Isolation, characterization and targeted disruption of mouse Ppia:
RL cyclophilin A is not essential for mammalian cell viability.";
DR EMBL; AF171073; AAD5096.1; -.
DR HSSP; P05092; 2CPL.
DR MGD; MGI:97749; Ppia.
DR InterPro; IPR002130; CSA_PPIase.
DR Pfam; PF00160; ppi_isomerase; 1.
DR PROSITE; PS50072; CSA_PPIase_2; 1.
FT NON TER
FT NON TER
SQ SEQUENCE 39 AA; 4324 MW; CBS3F70E1092889C CRC64;

Query Match 100.0%; Score 20; DB 11; Length 39;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
Db 26 GSSF 29

RESULT 7
C1888 PRELIMINARY; PRT; 39 AA.
ID C1888;
AC 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE YOSA protein.
GN YOSA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;

RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kuntz F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortlier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Hilbert H., Holsappel S., Hosono S., Hulech J., Harwood C.R., Henaut A.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Melillo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Priezen E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadala Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serot S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA Tosoato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winere P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis.";
```

```

RL Nature 390:249-256 (1997).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kuntz F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z99114; CAB13911.1; -.
KW Complete proteome.
SQ SEQUENCE 39 AA; 4090 MW; AC22034F15E873E1 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 39;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
Db 17 GSSF 20

RESULT 8
O9KV77 PRELIMINARY; PRT; 42 AA.
ID O9KV77;
AC O9KV77;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein VC0279.
GN VC0279.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_Taxid=666;

RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Dodeon R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bess S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uetebach T., Fleischmann R.D., Nielsen W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483 (2000).
DR EMBL; AE004116; AAF93454.1; -.
DR TIGR; VC0279; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 42 AA; 4964 MW; 1C3B4EBDD815BFEF CRC64;

Query Match 100.0%; Score 20; DB 16; Length 42;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
Db 38 GSSF 41

RESULT 9
O8X109 PRELIMINARY; PRT; 47 AA.
ID O8X109;
AC O8X109;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Laccase (EC 1.10.3.2) (Fragment).
GN LAC12.
OS Uncultured basidiomycete.
OC Eukaryota; Fungi; Basidiomycota; environmental samples.
OX NCBI_Taxid=175244;
RN (1)
RP SEQUENCE FROM N.A.
```

RC TISSUE-MYCELIUM;
 RA Luis P., Buscot F.;
 RT "molecular biological monitoring of soil fungi with laccase genes.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ420344; CAD12471.1; -;
 DR InterPro; IPR001117; Cu-oxidase.
 DR Pfam; PF00394; Cu-oxidase; 1.
 KW Oxidoreductase.
 FT NON_TER 1
 FT NON_TER 47
 SQ SEQUENCE 47 AA; 5419 MW; 3F5F74F8A8802097 CRC64;

Query Match 100.0%; Score 20; DB 3; Length 47;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 DB 27 GSSF 30

RESULT 10

Q9R4G2 PRELIMINARY; PRT; 20 AA.
 AC Q9R4G2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Beta-mannanase (EC 3.2.1.78) (fragment).
 OS Vibrio.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae.
 OX NCBI_TaxID=662;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96086028; PubMed=8534110;
 RA Tamari Y., Araki T., Amagoi H., Mori H., Morishita T.;
 RT "Purification and characterization of an extracellular beta-1,4-
 mannanase from a marine bacterium, *Vibrio* sp. strain MA-138.";
 RL Appl. Environ. Microbiol. 61:4454-4458(1995).
 SQ SEQUENCE 20 AA; 2192 MW; DB14359E0F4C7FC4 CRC64;

Query Match 85.0%; Score 17; DB 2; Length 20; —
 Best Local Similarity 75.0%; Pred. No. 9.2e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 DB 14 GSAP 17

RESULT 11

P82288 PRELIMINARY; PRT; 20 AA.
 AC P82288;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Glucan 1,6-beta-glucosidase (EC 3.2.1.-) (EXO-1,6-beta-
 (fragment).
 DE Acremonium sp.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Acremonium.
 OX NCBI_TaxID=5045;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=OXF C11;
 RA Jayus J., McDougall B.M., Seviour R.J.;
 RT "Purification and properties of a (1,6)-beta-glucanase from
 Acremonium sp. strain OXF C11."
 RL Submitted (JAN-2000) to the SWISS-PROT data bank.
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SUCCESSIVE GLUCOSE RESIDUES
 FROM 1,6-BETA-D-GLUCANS AND DERIVED OLIGOSACCHARIDES.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 KW Hydrolase; Glycosidase.

FT NON_TER 20
 SQ SEQUENCE 20 AA; 2175 MW; F831A01126E04B61 CRC64;
 Query Match 85.0%; Score 17; DB 3; Length 20;
 Best Local Similarity 75.0%; Pred. No. 9.2e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 DB 14 GSAP 17

RESULT 12

Q9TQY9 PRELIMINARY; PRT; 20 AA.
 AC Q9TQY9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 135 kDa glycoprotein/GP IIB homolog (fragment).
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96219641; PubMed=8639832;
 RA Thiabaudau K., Borche L., Soullilou J.P., Blanchard D.;
 RT "Characterization of porcine platelet glycoproteins recognized by
 human natural 'anti-gal' antibodies.";
 RL Blood 87:4636-4642(1996).
 SQ SEQUENCE 20 AA; 2092 MW; C2A46776631A33E CRC64;

Query Match 85.0%; Score 17; DB 6; Length 20;
 Best Local Similarity 75.0%; Pred. No. 9.2e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 DB 17 GSAP 20

RESULT 13

Q9R204 PRELIMINARY; PRT; 21 AA.
 AC Q9R204;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Heterogenous nuclear ribonucleoprotein A2/B1 (fragment).
 GN HNRNP A2/B1.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rosson M.J., DeGregori J., Ruley H.E.;
 RT "Gene trap mutagenesis of hnrnp A2/B1."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF073990; AAD29846.1; -;
 KW Nucleocapsid; Ribonucleoprotein.
 FT NON_TER 1
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 1937 MW; 940042C753673474 CRC64;

Query Match 85.0%; Score 17; DB 11; Length 21;
 Best Local Similarity 75.0%; Pred. No. 9.7e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSF 4
 DB 14 GSNF 17

RESULT 14

09WIE1 PRELIMINARY; PRT; 24 AA.
 AC 09WIE1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Defective RNA, strain T411.
 OS Citrus tristeza virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 CC Closterovirus.
 OX NCBI_TaxID=12162;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T411;
 RX MEDLINE=99190445; PubMed=10092024;
 RA Ayllon M.A., Lopez C., Navas-Castillo J., Mawassi M., Dawson W.O.,
 RA Guerri J., Flores R., Moreno P.;
 RT "New defective RNAs from citrus tristeza virus: evidence for a
 RT replicase-driven template switching mechanism in their generation.";
 RL J. Gen. Virol. 80:817-821(1999).
 DR EMBL; Y18368; CAA7716.1; -.
 SQ SEQUENCE 24 AA; 2691 MW; 277F9A071D51235A CRC64;

Query Match

Best Local Similarity 85.0%; Score 17; DB 12; Length 24;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
 DB 5 GSAF 8

RESULT 15

09EBR0 PRELIMINARY; PRT; 27 AA.
 AC 09EBR0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1027-3;
 RX MEDLINE=20411423; PubMed=10954550;
 RA Nelson J.A.E., Baribaud F., Edwards T., Swansstrom R.;
 RT "Patterns of Changes in Human Immunodeficiency Virus Type 1 V3
 RT Sequence Populations Late in Infection.";
 RL J. Virol. 74:8494-8501(2000).
 DR EMBL; AF155905; AAG09947.1; -.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 27 AA; 2874 MW; 8A5BE447659F47A7 CRC64;

Query Match

Best Local Similarity 85.0%; Score 17; DB 15; Length 27;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSF 4
 DB 13 GSAF 16

Search completed: January 10, 2003, 15:57:40
 Job time : 13.8182 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 10, 2003, 15:55:17 ; Search time 31.2727 Seconds
(without alignments)
34.087 Million cell updates/sec

Title: B
Perfect score: 40
Sequence: 1 gseflspe 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 433172

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneeq/geneeq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneeq/geneeq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneeq/geneeq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneeq/geneeq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneeq/geneeq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneeq/geneeq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneeq/geneeq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneeq/geneeq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneeq/geneeq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneeq/geneeq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneeq/geneeq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneeq/geneeq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneeq/geneeq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneeq/geneeq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneeq/geneeq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneeq/geneeq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneeq/geneeq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneeq/geneeq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneeq/geneeq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneeq/geneeq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneeq/geneeq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneeq/geneeq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneeq/geneeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	8	22	AAB60537 Ghrelin-like growt
2	40	100.0	9	22	AAB60538 Ghrelin-like growt
3	40	100.0	10	22	AAB60513 Ghrelin-like growt
4	40	100.0	11	22	AAB20100 SCIP peptide of 28
5	40	100.0	19	22	AAB62656 Human zslq33 pepti
6	40	100.0	27	22	AAB60514 Rat des-Gln14-ghre
7	40	100.0	27	22	AAB60515 Human des-Gln14-gn
8	40	100.0	27	22	AAB60519 Porcine des-Gln14-
9	40	100.0	27	22	AAB60522 Bovine ghrelin, SE
10	40	100.0	28	22	AAG64943 Neurone denaturati

11	40	100.0	28	22	AAB60508 Rat ghrelin, SEQ I
12	40	100.0	28	22	AAB60509 Human ghrelin, SEQ
13	40	100.0	28	22	AAB60518 Porcine ghrelin, S
14	40	100.0	28	22	AAB60530 Dog ghrelin-like G
15	40	100.0	28	23	AAE19032 Human ghrelin pept
16	40	100.0	28	23	AAE19041 Human ghrelin pept
17	36	90.0	8	23	AAU76320 Synthetic Ghrelin
18	36	90.0	10	23	AAE19026 Human ghrelin pept
19	36	90.0	14	23	AAU76321 Synthetic Ghrelin
20	36	90.0	14	23	AAE19022 Human ghrelin pept
21	36	90.0	14	23	AAE19023 Human ghrelin pept
22	36	90.0	18	23	AAE19025 Human ghrelin pept
23	36	90.0	23	23	AAE19024 Human ghrelin pept
24	36	90.0	28	22	AAB60560 Human ghrelin pept
25	36	90.0	28	23	AAE19021 Rat ghrelin-derive
26	36	90.0	28	23	AAE19027 Human ghrelin pept
27	36	90.0	28	23	AAE19028 Human ghrelin pept
28	36	90.0	28	23	AAE19029 Human ghrelin pept
29	36	90.0	28	23	AAE19030 Human ghrelin pept
30	36	90.0	28	23	AAE19031 Human ghrelin pept
31	36	90.0	28	23	AAE19033 Human ghrelin pept
32	36	90.0	28	23	AAE19034 Human ghrelin pept
33	36	90.0	28	23	AAE19035 Human ghrelin pept
34	36	90.0	28	23	AAE19036 Human ghrelin pept
35	36	90.0	28	23	AAE19037 Human ghrelin pept
36	36	90.0	28	23	AAE19038 Human ghrelin pept
37	36	90.0	28	23	AAE19039 Human ghrelin pept
38	36	90.0	28	23	AAE19040 Human ghrelin pept
39	35	87.5	7	22	AAB60507 Ghrelin-like growt
40	35	87.5	20	22	AAB60529 Rainbow trout 20aa
41	35	87.5	21	22	AAB60525 Bel ghrelin-like G
42	35	87.5	23	22	AAB60528 Rainbow trout 23aa
43	35	87.5	24	22	AAB60524 Chicken ghrelin-1i
44	30	75.0	33	22	AAB27736 Human peptide #187
45	30	75.0	33	22	ABB32907 Peptide #413 encod

ALIGNMENTS

PC515900/64907

RESULT 1	
AAAB60537	
ID	AAAB60537 standard; peptide; 8 AA.
XX	
AC	AAAB60537;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Ghrelin-like growth hormone secretagogue (GHS) core region peptide #3.
XX	
KW	Growth hormone secretagogue; GHS; ghrelin; core region;
KW	Calcium concentration elevation; infant growth disorder;
KW	growth hormone deficiency.
XX	
OS	Rattus norvegicus.
OS	Homo sapiens.
OS	Sus scrofa.
OS	Bos taurus.
XX	
PN	WO200107475-A1.
XX	
PD	01-FEB-2001. /
XX	
PF	24-JUL-2000; 2000MO-JP04907.
XX	
PR	23-JUL-1999; 99JP-0210002.
PR	29-NOV-1999; 99JP-0338841.
PR	26-APR-2000; 2000JP-0126623.
XX	
PA	(KANG/) KANGAWA K.
XX	
PI	Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y,
XX	

DR WPI; 2001-159704/16.
 XX New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 of infant growth disorders -
 XX
 PS Disclosure; Page 7; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
 XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 40; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8
 |||||
 DB 1 GSSFLSPE 8
 RESULT 2
 AAB60538
 ID AAB60538 standard; peptide; 9 AA.
 AC AAB60538;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Ghrelin-like growth hormone secretagogue (GHS) core region peptide #4.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin; core region;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Rattus norvegicus.
 OS Homo sapiens.
 OS Sus scrofa.
 OS Bos taurus.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 PA (KANG/) KANGAWA K.
 XX
 PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX
 DR WPI; 2001-159704/16.
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis

PT of infant growth disorders -
 XX
 PS Disclosure; Page 7; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 40; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8
 |||||
 DB 1 GSSFLSPE 8
 RESULT 3
 AAB60513
 ID AAB60513 standard; peptide; 10 AA. ✓
 AC AAB60513;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Ghrelin-like growth hormone secretagogue (GHS) core region, SEQ ID NO:9.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin; core region;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Rattus norvegicus.
 OS Homo sapiens.
 OS Sus scrofa.
 OS Bos taurus.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 PA (KANG/) KANGAWA K.
 XX
 PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX
 DR WPI; 2001-159704/16.
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 of infant growth disorders -
 XX
 PS Claim 3; Page 185; 210pp; Japanese.
 XX

CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptide. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-like growth hormone secretagogue (GHS) core region sequence.
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 40; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.097;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSFLSPE 8
 DB 1 GSSFLSPE 8
 RESULT 4
 AAB20100
 ID AAB20100 standard; Peptide; 11 AA. ✓
 AC AAB20100;
 XX 23-APR-2001 (first entry)
 DT
 DE SGIP peptide of zsig33.
 XX
 XX SGIP: zsig33; anorectic; antidiabetic; therapy; somatotropin;
 KM somatomedin-C; nutritional absorption modulator;
 KM growth hormone secretagogue; human.
 XX
 XX Homo sapiens.
 PH Key Location/Qualifiers
 FT Misc-difference 1 /note= "optionally substituted by Ser, Ala, Thr,
 FT Met in variant SGIP peptides"
 FT Misc-difference 2 /note= "optionally substituted by Gly, Ala, Thr,
 FT Met in variant SGIP peptides"
 FT Misc-difference 3 /note= "optionally substituted by Gly, Ala, Thr,
 FT Met in variant SGIP peptides"
 FT Misc-difference 4 /note= "optionally substituted by Trp, Tyr, Leu,
 FT Val, Ile in variant SGIP peptides"
 FT Misc-difference 5 /note= "optionally substituted by Ile, Val, Phe,
 FT Tyr in variant SGIP peptides"
 FT Misc-difference 6 /note= "optionally substituted by Gly, Ala, Thr,
 FT Met, Pro in variant SGIP peptides"
 FT Misc-difference 7 /note= "optionally substituted by Ala, Gly, Ile,
 FT Leu, Val in variant SGIP peptides"
 FT Misc-difference 8 /note= "optionally substituted by Asp in variant
 FT SGIP peptides"
 FT Misc-difference 9 /note= "optionally substituted by Arg, Lys, Phe,
 FT Tyr in variant SGIP peptides"
 FT Misc-difference 10

FT /note= "optionally substituted by Asn Ser, Thr,
 FT His, Ala, Glu, Asp, Lys, Arg in variant
 FT SGIP peptides"
 FT Misc-difference 11 /note= "optionally substituted by Gln, Asn, Ser,
 FT Thr, His, Ala in variant SGIP peptides"
 XX
 PN WO200100830-A1.
 XX
 PD 04-JAN-2001.
 XX
 XX 30-JUN-2000; 2000MO-US18306.
 XX
 XX 30-JUN-1999; 99US-0345157.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 PI
 XX WPI; 2001-123010/13.
 DR
 DR N-PSDB; AAF30032, AAF20034.
 XX
 PT Novel variants of SGIP peptides for modulating contractility in
 PT duodenum or jejunum tissue, pancreatic secretion of hormones and
 PT digestive enzymes, inducing growth hormone secretion or modulating
 PT gastric emptying -
 XX
 PS Claim 1: 53; 61pp; English.
 XX
 XX The present sequence is that of novel peptide fragment SGIP of
 CC zsig33 (see AAB20101), a previously described secreted protein
 CC that is transcribed in the gastrointestinal system. SGIP is a
 CC ligand for growth hormone secretagogue receptor, and is therefore
 CC useful for modulating secretion of growth hormone and insulin
 CC like growth factor 1. SGIP and its variant peptides, comprising
 CC residues 1-9, 2-9, 3-9, 4-9, 2-10, 3-10, 4-10, 3-11 or 4-11 of SGIP
 CC in which at least 1 residue may be substituted, are used in claimed
 CC methods for stimulating contractility in duodenum or jejunum
 CC tissue, modulating pancreatic secretion of hormones and digestive
 CC enzymes, inducing growth hormone secretion, and modulating gastric
 CC emptying. Amino acid substitutions in SGIP may result in agonist
 CC or antagonist activity. The SGIP peptides are also useful for
 CC raising antibodies, identifying the SGIP receptor, screening
 CC agonists and antagonists, and identifying cells, tissues or cell
 CC lines which respond to a SGIP-stimulated pathway.
 XX
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 40; DB 22; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSFLSPE 8
 DB 1 GSSFLSPE 8
 RESULT 5
 AAB62656
 ID AAB62656 standard; peptide; 19 AA. ✓
 AC AAB62656;
 XX 23-JUL-2001 (first entry)
 DT
 DE Human zsig33 peptide epitope hu3sig33-2.
 XX
 XX zsig33; signal transduction; hormone; enzyme; neural development;
 KM gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KM insulin-like growth factor-1; growth hormone; bone; gastrointestinal;
 KM glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
 KM G-protein coupled receptor.
 XX

OS Synthetic.
 OS Homo sapiens.
 PN WO200138355-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32074.
 XX
 PR 22-NOV-1999; 99US-0166765.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Sheppard PO, Jaapers SR, Deisher TA, Bishop PD;
 XX
 DR WPI; 2001-355879/37.
 XX
 XX Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX
 PS Example 8; Page 111; 11lpp; English.
 XX
 CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents a human zsig33
 CC peptide epitope, used to raise zsig33 anti-peptide antibodies.
 XX
 SQ Sequence 19 AA; —
 Query Match 100.0%; Score 40; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSFLSPE 8
 Db |||||
 1 GSSFLSPE 8
 RESULT 6
 AAB60514
 ID AAB60514 standard; peptide; 27 AA.
 AC AAB60514;
 XX
 XX 24-APR-2001 (first entry)
 DT
 DE Rat des-Gln14-ghrelin, SEQ ID NO:10.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200107475-A1.

XX PD 01-FEB-2001.
 XX
 XX 24-JUL-2000; 2000WO-JP04907.
 XX
 XX 23-JUL-1999; 99JP-0210002.
 PR 23-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 XX (KANG/) KANGAWA K.
 XX
 XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 PI WPI; 2001-159704/16.
 DR
 XX
 XX New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 XX
 PS Claim 3; Page 185; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 27 AA; —
 Query Match 100.0%; Score 40; DB 22; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSSFLSPE 8
 Db |||||
 1 GSSFLSPE 8
 RESULT 7
 AAB60515
 ID AAB60515 standard; peptide; 27 AA.
 XX
 AC AAB60515;
 XX
 XX 24-APR-2001 (first entry)
 DT
 DE Human des-Gln14-ghrelin, SEQ ID NO:11.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.

PR 26-APR-2000; 2000JP-0126623.
XX (KANG/) KANGAWA K.
XX
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
XX WPI; 2001-159704/16.
XX
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
XX Claim 3; Page 185; 210pp; Japanese.
XX
XX The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
XX
XX Sequence 27 AA:
SQ
Query Match 100.0%; Score 40; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSSFLSP 8
|||
1 GSSFLSP 8
DB 1 GSSFLSP 8
RESULT 8
AAB60519
ID AAB60519 standard; peptide; 27 AA. /
XX
XX AAB60519;
XX
XX 24-APR-2001 (first entry)
XX
XX Porcine des-Gln14-ghrelin, SEQ ID NO:17.
XX
XX Growth hormone secretagogue; GHS; ghrelin;
KW calcium concentration elevation; infant growth disorder;
KW growth hormone deficiency.
XX
XX Sus scrofa.
XX
XX WO200107475-A1.
XX
XX 01-FEB-2001.
XX
XX 24-JUL-2000; 2000WO-JP04907.
XX
XX 23-JUL-1999; 99JP-0210002.
PR 29-NOV-1999; 99JP-0338841.
PR 26-APR-2000; 2000JP-0126623.
XX
XX (KANG/) KANGAWA K.
XX
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
PI WPI; 2001-159704/16.
XX
XX

XX
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
XX Claim 3; Page 189; 210pp; Japanese.
XX
XX The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
XX
XX Sequence 27 AA:
SQ
Query Match 100.0%; Score 40; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSSFLSP 8
|||
1 GSSFLSP 8
DB 1 GSSFLSP 8
RESULT 9
AAB60522
ID AAB60522 standard; peptide; 27 AA. /
XX
XX AAB60522;
XX
XX 24-APR-2001 (first entry)
XX
XX Bovine ghrelin, SEQ ID NO:22.
XX
XX Growth hormone secretagogue; GHS; ghrelin;
KW calcium concentration elevation; infant growth disorder;
KW growth hormone deficiency.
XX
XX Bos taurus.
XX
XX WO200107475-A1.
XX
XX 01-FEB-2001.
XX
XX 24-JUL-2000; 2000WO-JP04907.
XX
XX 23-JUL-1999; 99JP-0210002.
PR 29-NOV-1999; 99JP-0338841.
PR 26-APR-2000; 2000JP-0126623.
XX
XX (KANG/) KANGAWA K.
XX
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
PI WPI; 2001-159704/16.
XX
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
XX Claim 3; Page 193; 210pp; Japanese.
XX
XX

CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 27 AA;

Query Match 100.0%; Score 40; DB 22; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8
 |||||
 DB 1 GSSFLSPE 8

RESULT 10
 AAG64943
 ID AAG64943 standard; peptide; 28 AA.

AC AAG64943;

DT 19-OCT-2001 (first entry)

DE Neurone denaturation prevention method related peptide #5.

XX Neurone denaturation; neurone death; growth hormone liberation inhibitor;
 KW cerebral infarction; oedema; Alzheimer's disease; Parkinson's disease;
 KW Pick's disease; dementia; amyotrophic lateral sclerosis; cancer;
 KW diabetic neuropathy; neuroprotective; antiinflammatory; nootropic;
 KW cytostatic.
 XX Unidentified.

XX Key Location/Qualifiers
 FH Modified-site 3 /label= OTHER
 FT /note= "modified by O(C=O) (CH2)6CH3"
 FT

XX WO200147558-A1.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09431.

XX 28-DEC-1999; 99JP-0375513.

XX (KAKE) KAKEN PHARM CO LTD.

XX Murata T, Ohyama T, Amakawa M, Fujita K, Ueo H;

XX WPI; 2001-536280/59.

XX Agents for treating diseases associated with denaturation or death of
 PT neurons comprise growth hormone liberation inhibitor -

XX Disclosure; Page 17; 50pp; Japanese.

XX The present invention provides agents for treating or preventing diseases
 CC associated with denaturation or death of neurons, which comprise a
 CC growth hormone liberation inhibitor. These can be used for treating or

CC preventing diseases associated with denaturation or death of neurons
 CC including those due to cerebral ischaemic disorders such as cerebral
 CC infarction or oedema. Other causes of denaturation or death of neurons
 CC included Alzheimer's disease, Pick's disease, AIDS related dementia,
 CC Parkinson's disease, amyotrophic lateral sclerosis, diabetic neuropathy
 CC and anticancer treatments. The present sequence is a peptide described in
 CC the exemplification of the invention.
 XX

SQ Sequence 28 AA;

Query Match 100.0%; Score 40; DB 22; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSFLSPE 8
 |||||
 DB 1 GSSFLSPE 8

RESULT 11
 AAB60508
 ID AAB60508 standard; peptide; 28 AA.

XX AAB60508;

DT 24-APR-2001 (first entry)

DE Rat ghrelin, SEQ ID NO:2.

XX Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.

XX Rattus norvegicus.

XX WO200107475-A1.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-JP04907.

XX 23-JUL-1999; 99JP-0210002.

XX 29-NOV-1999; 99JP-0338841.

XX 26-APR-2000; 2000JP-0126623.

XX (KANG/) KANGAWA K.

XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;

XX WPI; 2001-159704/16.

XX New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -

XX Claim 2; Page 180; 210pp; Japanese.

XX The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.

XX Sequence 28 AA;
SQ

Query Match
Best Local Similarity 100.0%; Score 40; DB 22; Length 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSFLSP 8
DB 1 GSSFLSP 8

RESULT 12
AAB60509
ID AAB60509 standard; peptide; 28 AA.

XX AAB60509;
XX
XX 24-APR-2001 (first entry)
XX
XX Human ghrelin, SEQ ID NO:3.
XX
XX Growth hormone secretagogue; GHS; ghrelin;
XX
XX calcium concentration elevation; infant growth disorder;
XX
XX growth hormone deficiency.
XX
XX Homo sapiens.
XX
XX WO200107475-A1.
XX
XX 01-FEB-2001.
XX
XX 24-JUL-2000; 2000MO-JP04907.
XX
XX 23-JUL-1999; 99JP-0210002.
XX
XX 29-NOV-1999; 99JP-0338841.
XX
XX 26-APR-2000; 2000JP-0126623.
XX
XX (KANG/) KANGAWA K.
XX
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
XX
XX WPI; 2001-159704/16.
XX
XX
XX New peptide compounds which induce growth hormone secretion and
XX
XX elevate cell calcium concentrations, useful in treatment and diagnosis
XX
XX of infant growth disorders -
XX
XX
XX Claim 3; Page 181; 210pp; Japanese.

CC The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-type growth hormone secretagogue (GHS) of the invention.

SQ Sequence 28 AA;

Query Match
Best Local Similarity 100.0%; Score 40; DB 22; Length 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSFLSP 8
DB 1 GSSFLSP 8

RESULT 13
AAB60518
ID AAB60518 standard; peptide; 28 AA.

XX AAB60518;
XX
XX 24-APR-2001 (first entry)
XX
XX Porcine ghrelin, SEQ ID NO:16.
XX
XX
XX Growth hormone secretagogue; GHS; ghrelin;
XX
XX calcium concentration elevation; infant growth disorder;
XX
XX growth hormone deficiency.
XX
XX Sus scrofa.
XX
XX WO200107475-A1.
XX
XX 01-FEB-2001.
XX
XX 24-JUL-2000; 2000MO-JP04907.
XX
XX 23-JUL-1999; 99JP-0210002.
XX
XX 29-NOV-1999; 99JP-0338841.
XX
XX 26-APR-2000; 2000JP-0126623.
XX
XX (KANG/) KANGAWA K.
XX
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
XX
XX WPI; 2001-159704/16.
XX
XX
XX New peptide compounds which induce growth hormone secretion and
XX
XX elevate cell calcium concentrations, useful in treatment and diagnosis
XX
XX of infant growth disorders -
XX
XX
XX Claim 3; Page 189; 210pp; Japanese.

CC The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-type growth hormone secretagogue (GHS) of the invention.

SQ Sequence 28 AA;

Query Match
Best Local Similarity 100.0%; Score 40; DB 22; Length 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSSFLSP 8
DB 1 GSSFLSP 8

RESULT 14
AAB60530

ID AAB60530 standard; peptide; 28 AA.
 XX AAB60530;
 AC 24-APR-2001 (first entry)
 DT Dog ghrelin-like GH secretagogue peptide, SEQ ID NO:31.
 DE Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX Canis familiaris.
 OS WO200107475-A1.
 PN 01-FEB-2001.
 XX 24-JUL-2000; 2000WO-JP04907.
 XX 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX (KANG/) KANGAWA K.
 PA Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 PI WPI; 2001-159704/16.
 DR New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 PS Claim 4; Page 197; 210pp; Japanese.
 XX The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX Sequence 28 AA;
 SQ
 Query Match 100.0%; Score 40; DB 22; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GSSFSLSP 8
 Db 1 GSSFSLSP 8
 RESULT 15
 ID AAE19032 standard; peptide; 28 AA.
 XX AAE19032;
 AC 21-MAY-2002 (first entry)
 DT Human ghrelin peptide analogue, compound 6.
 DE

XX Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
 KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
 KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
 KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
 KW cardiovascular disorder; gall stones; osteoarthritis; cancer; cytostatic;
 KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;
 KW cardiant; litholytic; hepatotropic.
 XX Homo sapiens.
 OS WO200192292-A2.
 PN 06-DEC-2001.
 XX 25-MAY-2001; 2001WO-US17026.
 XX 30-MAY-2000; 2000US-207920P.
 XX (MERI) MERCK & CO INC.
 PA Bednarek M;
 PI WPI; 2002-195531/25.
 DR Truncated ghrelin analogs active at growth-hormone secretagogue
 PT receptor useful for diagnosing or treating diseases such as anorexia,
 PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
 PT osteoarthritis -
 XX Example 4; Page 34; 37pp; English.
 XX The present invention relates to a truncated ghrelin analogue or their
 CC salt, active at growth-hormone secretagogue (GHS) receptor. Ghrelin
 CC analogue is useful for screening a compound capable of binding to GHS
 CC receptor and for stimulating growth hormone secretion. Ghrelin agonist
 CC is utilised for treating a growth hormone deficient state, increasing
 CC muscle mass and bone density, treating sexual dysfunction in males or
 CC females, facilitating a weight gain, maintenance of weight, maintenance
 CC of physical functioning, recovery of physical function, and/or appetite
 CC increase, or appetite increase is particularly useful for a patient
 CC having a disease or disorder, or under going a treatment, accompanied by
 CC eight less such as anorexia, bulimia, cancer cachexia, acquired
 CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
 CC elderly and examples of treatments accompanied by weight loss include
 CC chemotherapy, radiation therapy, temporary or permanent immobilisation
 CC and dialysis; and ghrelin antagonist is utilised to facilitate weight
 CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
 CC complications of diabetes including retinopathy, and/or cardiovascular
 CC disorders, where excessive weight is a contributing factor to different
 CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
 CC disease, gall stones, osteoarthritis and certain forms of cancers, and
 CC bringing about a weight loss can be used for e.g. to reduce the
 CC likelihood of such diseases and for treating such diseases. Ghrelin
 CC analogue induces growth hormone release from primary-culture pituitary
 CC cells in a dose-dependent manner without stimulating the release of other
 CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
 CC synthesised easily and has increased solubility in physiological buffers.
 CC The present sequence is human ghrelin peptide analogue.
 XX Sequence 28 AA;
 SQ
 Query Match 100.0%; Score 40; DB 23; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GSSFSLSP 8
 Db 1 GSSFSLSP 8
 Search completed: January 10, 2003, 15:59:13
 Job time : 31.2727 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 10, 2003, 15:55:16 ; Search time 12.7273 seconds
(without alignments)
60.427 Million cell updates/sec

Title: B
Perfect score: 40
Sequence: 1 gseflspe 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 11827

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	72.5	25	2	S03275
2	26	65.0	27	2	PC2337
3	26	65.0	27	2	PC2338
4	26	65.0	33	2	B44644
5	25	62.5	30	2	A28835
6	24	60.0	20	2	T26748
7	24	60.0	42	2	S20329
8	23	57.5	47	2	PN0607
9	23	57.5	48	2	S01122
10	22	55.0	12	2	S01122
11	22	55.0	15	2	PA0027
12	22	55.0	21	2	I49414
13	22	55.0	37	2	A81552
14	22	55.0	38	2	S04627
15	22	55.0	43	2	F83732
16	22	55.0	44	2	T07534
17	22	55.0	47	2	AF3178
18	22	55.0	48	2	I38223
19	22	55.0	50	2	A70242
20	21	52.5	17	2	A37823
21	21	52.5	19	2	S69166
22	21	52.5	22	2	C60691
23	21	52.5	24	2	J73619
24	21	52.5	35	2	J70519
25	21	52.5	38	2	S78757
26	21	52.5	39	2	I40555
27	21	52.5	39	2	A82707
28	21	52.5	39	2	T12912
29	21	52.5	40	2	A19940

30	21	52.5	40	2	S71301	ICU5 protein - Par
31	21	52.5	40	2	B97413	hypothetical prote
32	21	52.5	41	2	T12846	hypothetical prote
33	21	52.5	42	2	S41210	serine proteinase
34	21	52.5	42	2	B82657	hypothetical prote
35	21	52.5	50	2	H90760	hypothetical prote
36	21	52.5	50	2	G81239	hypothetical prote
37	21	52.5	50	2	G97539	hypothetical prote
38	20	50.0	18	2	S23950	45k protein - pig
39	20	50.0	20	2	A25335	myosin-light-chain
40	20	50.0	21	2	S46550	actin-related prot
41	20	50.0	25	2	PH1733	Ig heavy chain V r
42	20	50.0	27	2	PC2339	second envelope g1
43	20	50.0	28	2	PC4429	peroxisome prolif
44	20	50.0	30	2	S66439	allophycocyanin 11
45	20	50.0	33	2	PH1738	Ig heavy chain V r

ALIGNMENTS

RESULT 1
S03275
photosystem II 22K protein - spinach (fragment)
C:Species: Spinacia oleracea (spinach)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Aug-1998
C:Accession: S03275
R:Murata, N.; Kajimura, H.; Fujimura, Y.; Miyao, M.; Murata, T.; Watanabe, A.; Shinozaki, Prog. Photosyn. Res. 1, 701-704, 1987
A:Title: Partial amino acid sequences of the proteins of pea and spinach photosystem II
A:Reference number: S03269
A:Accession: S03275
A:Molecule type: protein
A:Residues: 1-25 <MUR>
C:Superfamily: chlorophyll a/b-binding protein
C:Keywords: chloroplast; photosynthesis; photosystem II

Query Match
Best Local Similarity 72.5%; Score 29; DB 2; Length 25;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGSFLSPE 8
Db 16 GGNFLDPE 23

RESULT 2
PC2337
second envelope glycoprotein, gp70, hypervariable region 1 - hepatitis C virus (isolate I
C:Species: hepatitis C virus
C:Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C:Accession: PC2337
R:Kato, N.; Nakazawa, T.; Mizutani, T.; Shimotohno, K. Blochem. Biophys. Res. Commun. 206, 863-869, 1995
A:Title: Susceptibility of human T-lymphotropic virus type I infected cell line MT-2 to
A:Reference number: PC2334; MUID:95134269; PMID:7832798
A:Accession: PC2337
A:Molecule type: genomic RNA
A:Residues: 1-27 <KAT>
A:Cross-references: DBJ:D43649; MID:9882083; PIDN:BA07764.1; PID:d1008349; PID:9882084
A:Experimental source: isolate B-1, inoculum, 1 day and 4 days postinoculation cells
C:Keywords: glycoprotein

Query Match
Best Local Similarity 65.0%; Score 26; DB 2; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSFLSP 7
Db 17 TSFLSP 22

RESULT 3

PC2338
second envelope glycoprotein, gp70, hypervariable region 1 - hepatitis C virus (isolate
C;Species: hepatitis C virus
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C;Accession: PC2338
R;Kato, N.; Nakazawa, T.; Mizutani, T.; Shimotohno, K.
Biochem. Biophys. Res. Commun. 206, 863-869, 1995
A;Title: Susceptibility of human T-lymphotropic virus type I infected cell line MT-2 to
A;Reference number: PC2334; MUID:95134269; PMID:7832798
A;Accession: PC2338
A;Molecule type: genomic RNA
A;Residues: 1-27 <KAT>
A;Cross-references: DDBJ:D43650; NID:G882085; PIDN:BAA07765.1; PID:dl008350; PID:G882086
A;Experimental source: isolate B-2, inoculum
C;Keywords: glycoprotein

Query Match 65.0%; Score 26; DB 2; Length 27;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 SSFLSP 7
Db 17 TSFLSP 22

RESULT 4
B44644
neurotoxin-associated protein type B Hn+ 17K chain - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
C;Accession: B44644
R;Somers, E.; DasGupta, B.R.
J. Protein Chem. 10, 415-425, 1991
A;Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without he
A;Reference number: A44644; MUID:92143938; PMID:1781887
A;Contents: type B
A;Accession: B44644
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-33 <SOM>
A;Note: sequence extracted from NCBI backbone (NCBIP:83799)
A;Note: 16-Val was also found
C;Keywords: hemagglutinin

Query Match 65.0%; Score 26; DB 2; Length 33;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSFLSP 7
Db 20 GSYLSP 26

RESULT 5
A28835
Ig heavy chain V region (3E11) - channel catfish (fragment)
C;Species: Ictalurus punctatus (channel catfish)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Aug-1996
C;Accession: A28835
R;Lobb, C.J.; Olson, M.O.J.
J. Immunol. 141, 1236-1245, 1988
A;Title: Immunoglobulin heavy H chain isotypes in a teleost fish.
A;Reference number: A92822; MUID:98285736; PMID:2456346
A;Accession: A28835
A;Molecule type: protein
A;Residues: 1-30 <LOB>
C;Keywords: heterotetramer; immunoglobulin

Query Match 62.5%; Score 25; DB 2; Length 30;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSFLSP 7

Db 7 GSTVLSLP 13
||: |||

RESULT 6

T26748
hypothetical protein Y39A1B.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26748
R;Wall, M.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20258
A;Accession: T26748
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-20 <WIL>
A;Cross-references: EMBL:AL021482; PIDN:CAB54436.1; GSPDB:GN00021; CESP:Y39A1B.4
A;Experimental source: clone Y39A1B
C;Genetics:
A;Gene: CESP:Y39A1B.4
A;Map position: 3
A;Introns: 16/3

Query Match 60.0%; Score 24; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 SFLSPE 8
Db 14 SFLVPE 19

RESULT 7

S20329
beta-1,3-glucanase - kidney bean
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 13-Sep-1998
C;Accession: S20329
R;Daugrois, J.H.; Lafitte, C.; Barthe, J.P.; Faucher, C.; Touze, A.; Esquerre-Tugaye, M.
Arch. Biochem. Biophys. 292, 468-474, 1992
A;Title: Purification and characterization of two basic beta-1,3-glucanases induced in
A;Reference number: S20329; MUID:92117656; PMID:1731612
A;Accession: S20329
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-42 <DAU>
C;Superfamily: beta-1,3-glucanase

Query Match 60.0%; Score 24; DB 2; Length 42;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GSSFLSP 7
Db 26 GSEYLDLP 32

RESULT 8

PN0607
cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: PN0607
R;Sillard, R.; Joenvall, H.; Mutt, V.
Biochem. Biophys. Res. Commun. 195, 746-750, 1993
A;Title: Characterization of porcine intestinal cytochrome c oxidase subunit VIIC, puri
A;Reference number: PN0607; MUID:93384597; PMID:8396926
A;Accession: PN0607
A;Molecule type: protein
A;Residues: 1-47 <SIL>
A;Experimental source: intestine
C;Superfamily: cytochrome-c oxidase chain VIIC

C:Keywords: mitochondrion; oxidoreductase; respiratory chain

Query Match 57.5%; Score 23; DB 2; Length 47;

Best Local Similarity 57.1%; Pred. No. 5.1e+02; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFLSP 7
| | | | |
DB 30 GSCFAP 36

RESULT 9

138221

protein-serine/threonine kinase - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 24-Sep-1999

C:Accession: 138221; S37423

R:Schultz, S.J.; Ni99, E.A.

Cell Growth Differ. 4, 821-830, 1993

A:Title: Identification of 21 novel human protein kinases, including 3 members of a family

A:Reference number: 138211; MUID:94100173; PMID:8274451

A:Accession: 138221

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-48 <RES>

A:Cross-references: EMBL:225431; NID:9405740; PIDN:CAA80918.1; PID:9405741

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 57.5%; Score 23; DB 2; Length 48;

Best Local Similarity 80.0%; Pred. No. 5.2e+02; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLSP 8
| | | | |
DB 27 YLSP 31

RESULT 10

S01122

photosystem II 3.7k protein - spinach (fragment)

C:Species: Spinacia oleracea (spinach)

C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 18-Jun-1993

C:Accession: S01122

R:Schroeder, W.P.; Henryson, T.; Akerlund, H.B.

FEBS Lett. 235, 289-297, 1988

A:Title: Characterization of low molecular mass proteins of photosystem II by N-terminal

A:Reference number: S01120

A:Molecule type: protein

A:Residues: 1-12 <SCH>

A:Keywords: chloroplast; photosynthesis; photosystem II; thylakoid

Query Match 55.0%; Score 22; DB 2; Length 12;

Best Local Similarity 80.0%; Pred. No. 1.8e+02; Mismatches 1; Indels 0; Gaps 0;

QY 3 SFUSP 7
| | | | |
DB 7 AFUSP 11

RESULT 11

PA0027

protein QA10006 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997

C:Accession: PA0027

R:Kamo, M.; Kawakami, T.; Miyake, N.; Tsugita, A.

submitted to JIPD, July 1994

A:Description: Separation and characterization of Arabidopsis proteins by two-dimensions

A:Reference number: PA0001

A:Accession: PA0027

A:Molecule type: protein

A:Residues: 1-15 <RAM>

A:Experimental source: callus

Query Match 55.0%; Score 22; DB 2; Length 15;

Best Local Similarity 50.0%; Pred. No. 2.3e+02; Mismatches 3; Indels 0; Gaps 0;

QY 1 GSSFLSP 8
| | | | |
DB 6 GPKFASPK 13

RESULT 12

I49414

gene CTLA-1 protein - western wild mouse (fragment)

C:Species: Mus spretus (western wild mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999

C:Accession: I49414

R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.

Mamm. Genome 5, 349-355, 1994

A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.

A:Reference number: I48934; MUID:94319082; PMID:8043949

A:Accession: I49414

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-21 <RES>

A:Cross-references: EMBL:U05708; NID:9497037; PIDN:AAB60471.1; PID:9497038

C:Superfamily: trypsin; trypsin homology

Query Match 55.0%; Score 22; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 3.4e+02; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLS 6
| | | | |
DB 8 SSFLS 12

RESULT 13

A81552

hypothetical protein CP0663 [imported] - Chlamydia pneumoniae (strain AR39)

C:Species: Chlamydia pneumoniae

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000

C:Accession: A81552

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, J.

C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: A81552

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-37 <RBA>

A:Cross-references: GB:AE002224; GB:AE002161; NID:97189574; PIDN:AAF38475.1; PID:97189577.

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CP0663

Query Match 55.0%; Score 22; DB 2; Length 37;

Best Local Similarity 66.7%; Pred. No. 6.4e+02; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSFLSP 7
| | | | |
DB 29 SDFLTP 34

RESULT 14

S04627

glutathione transferase (EC 2.5.1.18) 6.0 - Proteus mirabilis (fragment)

C:Species: Proteus mirabilis

C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 01-Feb-1999
 C:Accession: S04627
 R:Di Ilio, C.; Aceto, A.; Piccolomini, R.; Allocati, N.; Caccuri, A.M.; Barra, D.; Fedez
 FEBS Lett. 250, 57-59, 1989
 A:Title: N-terminal region of Proteus mirabilis glutathione transferase is not homologous
 A:Reference number: S04627; MUID:89290034; PMID:2661269

A:Accession: S04627
 A:Molecule type: protein
 A:Residues: 1-38 <DI>
 C:Complex: dimer

C:Function:

A:Description: catalyzes conjugation of glutathione to a large variety of electrophilic
 bic compounds; involved in detoxification of organic hydroperoxides
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; transferase

Query Match 55.0%; Score 22; DB 2; Length 38;
 Best Local Similarity 71.4%; Pred. No. 6.6e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFLSP 7
 ||||
 Db 8 GSCSLSP 14

RESULT 15

F83732
 hypothetical protein BH0662 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: F83732
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, P.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: F83732
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-43 <STO>
 A:Cross-references: GB:AF001509; GB:BA000004; NID:gl0173176; PIDN:BA04381.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0662

Query Match 55.0%; Score 22; DB 2; Length 43;
 Best Local Similarity 66.7%; Pred. No. 7.6e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SFLSP 8
 ||||
 Db 16 SFVPP 21

Search completed: January 10, 2003, 15:56:29
 Job time : 14.7273 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:16 ; Search time 7.27273 Seconds
(without alignments)
45.624 Million cell updates/sec

Title: B
Perfect score: 40
Sequence: 1 gseflspe 8

Scoring table: BLOSSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 3754

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	60.0	46	1	CSPA_KLEPN Q48493 klebsiella
2	23	57.5	46	1	CSPA_AERHY Q44078 aeromonas h
3	23	57.5	46	1	CSPA_AERSA Q44317 aeromonas s
4	22	55.0	46	1	CSPA_CITPR Q46051 citrobacter
5	22	55.0	46	1	CSPA_ENTAE Q46664 enterobacte
6	22	55.0	46	1	CSPA_SALVI Q56178 salmonella
7	22	55.0	46	1	CSPA_SHIBO Q53816 shigella bo
8	22	55.0	46	1	CSPA_SHIFL Q54170 shigella fl
9	21	52.5	15	1	FKB7_PINPS P81104 pinus pins
10	21	52.5	46	1	GP54_BRSPL Q48408 bacterioph
11	20	50.0	19	1	OXLA_ORHHA P81383 ophiophagu
12	20	50.0	30	1	TLIX_SPTOL P82337 spiniactia ol
13	20	50.0	32	1	IAPB_BOVIN Q28207 bos taurus
14	20	50.0	32	1	IAPB_SHEEP Q28605 ovie aries
15	20	50.0	37	1	ATPO_SOLTU P80504 solanum tub
16	20	50.0	41	1	LH83_RHOAC P35096 rhodospende
17	20	50.0	45	1	CSPA_STRDY Q53984 streptococ
18	20	50.0	46	1	CSPA_YAREN Q56922 yersinia en
19	20	50.0	49	1	R22A_MOUSE P35385 mus musculu
20	20	50.0	50	1	R22A_MOUSE P82911 bos taurus
21	19	47.5	27	1	BLP2_BOMOR P29003 bomina ori
22	19	47.5	31	1	GT_SERMA P22416 serratia ma
23	19	47.5	36	1	TXJB_HADVE P82227 hadronyche
24	19	47.5	36	1	TXJB_HADVE P82226 hadronyche
25	19	47.5	39	1	DHPS_ERWCI P19574 erwinia cit
26	19	47.5	39	1	PSBK_PORPU P51197 porphyra pu
27	19	47.5	40	1	HC3S_THIFE P80510 thibacillu
28	19	47.5	40	1	NLTB_WHEAT P26013 triticum ae
29	19	47.5	43	1	PSBK_CYARA P48108 cyanophora
30	19	47.5	46	1	PSAI_ANASP P58560 anabaena sp
31	19	47.5	46	1	PSAI_ANAVA P23079 anabaena va
32	19	47.5	46	1	PSBK_PEA P28642 pisum sativ
33	19	47.5	47	1	VARP_BPP22 P14112 bacterioph

34	19	47.5	47	1	VNST_BUNL7 P09614 bunyavirus
35	19	47.5	48	1	ATP8_HANNI P48882 hanseniella w
36	19	47.5	48	1	YO48_BPHKO Q37928 bacterioph
37	19	47.5	48	1	YO48_BPP22 Q36661 bacterioph
38	19	47.5	49	1	RBSG_MOUSE P35378 mus musculu
39	19	47.5	49	1	Y4ZD_RHISN P55731 rhizobium s
40	18	45.0	8	1	LCKS_LEUMA P19987 leucophaea
41	18	45.0	12	1	CXL3_CONNR P58809 conus marmo
42	18	45.0	20	1	JHPB_BOMMO P81627 bombyx mori
43	18	45.0	21	1	LPRM_CORDI P21232 corynebacte
44	18	45.0	21	1	SCIB_BPTS P23208 bacterioph
45	18	45.0	25	1	POTX_PARCV P41736 paraponera

ALIGNMENTS

```

RESULT 1
CSPA_KLEPN STANDARD; PRT: 46 AA.
ID CSPA_KLEPN
AC Q48493:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 9633;
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;
RT "Detection and speculation of bacteria through PCR using universal
RT major cold-shock protein primer oligomers."
RL J. Ind. Microbiol. Biotechnol. 19:286-293 (1997).
CC - SUBUNIT: HOMODIMER (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC - INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
EMBL: U60039; AAC80243.1; -.
DR HSSP: P15277; IMTC.
DR InterPro: IPR002059; Cold_shock.
DR Pfam: PF00313; CSD; 1.
DR ProDom: PD000621; Cold_shock; 1.
DR SMART: SM00357; CSD; 1.
DR PROSITE: PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON TER 1 >46 CSD.
FT DOMAIN 1 >46 CSD.
FT NON TER 46
SQ SEQUENCE 46 AA; 4982 MW; 17031DD393760DEC CRC64;
Query Match 60.0%; Score 24; DB 1; Length 46;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GSEFLSPE 8
DB 3 GSEFLSPE 10

```

```

RESULT 2
CSPA_AERHY
ID_CSPA_AERHY STANDARD; PRT; 46 AA.
AC Q44078;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPI.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;
RT "Detection and speciation of bacteria through PCR using universal
major cold-shock protein primer oligomers.";
RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60026; AAC80230.1; -
DR HSP; P15277; IMJC.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON_TER 1 1
FT DOMAIN <1 >46 CSD.
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5105 MW; 3AF60F50D0C30D41 CRC64;

Query Match 57.5%; Score 23; DB 1; Length 46;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFLSP 7
| | | | |
Db 3 GFGFISP 9

RESULT 3
CSPA_AERSA
ID_CSPA_AERSA STANDARD; PRT; 46 AA.
AC Q44317;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPI.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIMB 1102;
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;

```

```

RT "Detection and speciation of bacteria through PCR using universal
major cold-shock protein primer oligomers.";
RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60027; AAC80231.1; -
DR HSP; P15277; IMJC.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON_TER 1 1
FT DOMAIN <1 >46 CSD.
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5095 MW; 3AF60F4A10DD1661 CRC64;

Query Match 57.5%; Score 23; DB 1; Length 46;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFLSP 7
| | | | |
Db 3 GFGFISP 9

RESULT 4
CSPA_CITFR
ID_CSPA_CITFR STANDARD; PRT; 46 AA.
AC Q46051; Q45969;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPI.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCTC 6071, and NCTC 6081 / C.ballerapensis;
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;
RT "Detection and speciation of bacteria through PCR using universal
major cold-shock protein primer oligomers.";
RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60032; AAC80236.1; -

```

```

DR EMBL; U60033; AAC00237.1; -.
DR HSSP; P15277; IMTC.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON TER 1 1
FT DOMAIN <1 >46 CSD.
FT NON TER 46 46
SQ SEQUENCE 46 AA; 4983 MW; 5B38560260B9E935 CRC64;

Query Match
Best Local Similarity 55.0%; Score 22; DB 1; Length 46;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GSSFLSPE 8
Db 3 GFGFIRPD 10

RESULT 5
CSPA_ENTAE STANDARD; PRT; 46 AA.
AC Q46664;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPA.
OS Enterobacter aerogenes (Aerobacter aerogenes).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=548;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCITC 10006;
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;
RT "Detection and speciation of bacteria through PCR using universal
RT major cold-shock protein primer oligomers.";
RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U60034; AAC00238.1; -.
DR HSSP; P15277; IMTC.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON TER 1 1
FT DOMAIN <1 >46 CSD.
FT NON TER 46 46
SQ SEQUENCE 46 AA; 4983 MW; 5B38560260B9E935 CRC64;

Query Match
Best Local Similarity 55.0%; Score 22; DB 1; Length 46;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy 1 GSSFLSPE 8
Db 3 GFGFIRPD 10

RESULT 6
CSPA_SALVI STANDARD; PRT; 46 AA.
AC Q56178;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPA.
OS Salmonella virchow.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=48409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCITC 5742;
RX MEDLINE=98101990; PubMed=9439003;
RA Francis K.P., Stewart G.S.A.B.;
RT "Detection and speciation of bacteria through PCR using universal
RT major cold-shock protein primer oligomers.";
RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U60049; AAC00253.1; -.
DR HSSP; P15277; IMTC.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator.
FT NON TER 1 1
FT DOMAIN <1 >46 CSD.
FT NON TER 46 46
SQ SEQUENCE 46 AA; 4983 MW; 5B38560260B9E935 CRC64;

Query Match
Best Local Similarity 55.0%; Score 22; DB 1; Length 46;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GSSFLSPE 8
Db 3 GFGFIRPD 10

RESULT 7
CSPA_SHIBO STANDARD; PRT; 46 AA.
AC Q53816;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major cold-shock protein (Fragment).
GN CSPA.
OS Shigella boydii.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.

```

OX NCBI_TaxID=621;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 9327;
 RX MEDLINE=98101990; PubMed=9439003;
 RA Francis K.P., Stewart G.S.A.B.;
 RT "Detection and speciation of bacteria through PCR using universal
 RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U60036; AAC80240.1; -;
 DR HSSP; P15277; IMJC.
 DR InterPro; IPR002059; Cold_shock.
 DR Pfam; PF00313; CSD; 1.
 DR ProDom; PD000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1; Shock; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 KW Transcription regulation; DNA-binding; Activator.
 FT NON_TER 1 1
 FT DOMAIN <1 >46 CSD.
 FT NON_TER 46 46
 FT SEQUENCE 46 AA; 4983 MW; 5B38560260B9E935 CRC64;
 SQ
 Query Match 55.0%; Score 22; DB 1; Length 46;
 Best Local Similarity 37.5%; Pred. NO. 3.5e+02;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GSSFLSPE 8
 DB 3 GFGFTPD 10
 RESULT 8
 CSQA SHIFL STANDARD; PRT; 46 AA.
 AC Q54170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major cold-shock protein (Fragment).
 GN CSQA.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 0002;
 RX MEDLINE=98101990; PubMed=9439003;
 RA Francis K.P., Stewart G.S.A.B.;
 RT "Detection and speciation of bacteria through PCR using universal
 RL J. Ind. Microbiol. Biotechnol. 19:286-293(1997).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- INDUCTION: IN RESPONSE TO LOW TEMPERATURE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U60037; AAC80241.1; -;
 DR HSSP; P15277; IMJC.
 DR InterPro; IPR002059; Cold_shock.
 DR Pfam; PF00313; CSD; 1.
 DR ProDom; PD000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1; Shock; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 KW Transcription regulation; DNA-binding; Activator.
 FT NON_TER 1 1
 FT DOMAIN <1 >46 CSD.
 FT NON_TER 46 46
 FT SEQUENCE 46 AA; 4983 MW; 5B38560260B9E935 CRC64;
 SQ
 Query Match 55.0%; Score 22; DB 1; Length 46;
 Best Local Similarity 37.5%; Pred. NO. 3.5e+02;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GSSFLSPE 8
 DB 3 GFGFTPD 10
 RESULT 9
 FK7 PINPS STANDARD; PRT; 15 AA.
 AC P81104;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE 70 kDa peptidylprolyl isomerase (EC 5.2.1.8) (Peptidylprolyl cis-trans
 DE isomerase) (Cyclophilin) (Ppiase) (S1205-06) (Fragment).
 OS Pinus pinaster (Maritime pine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=71647;
 RN [1]
 RP SEQUENCE.
 RC Tissue=Needle;
 RA Plomion C., Costa P., Bahrman N., Frigerio J.-M.;
 RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
 RT dominant and codominant protein markers assayed on diploid tissue, in
 RT a haploid-based genetic map.";
 RL Silvae Genetica 46:161-165(1997).
 RN [2]
 RP SEQUENCE.
 RC Tissue=Needle;
 RX MEDLINE=99274088; PubMed=10344291;
 RA Costa P., Plomion C., Bauw G., Dubos C., Bahrman N., Kremer A.,
 RA Frigerio J.-M., Plomion C.;
 RT "Separation and characterization of needle and xylem maritime pine
 RT proteins.";
 RL Electrophoresis 20:1098-1108(1999).
 CC -1- FUNCTION: PPIASE THAT BINDS CALMODULIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
 CC PEPTIDE BONDS IN OLIGOPEPTIDES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
 CC 5-3. ITS MW IS: 72 kDa.
 CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
 DR InterPro; IPR001179; FKBP_PPIase.
 DR PROSITE; PS00453; FKBP_PPIASE_1; PARTIAL.
 DR PROSITE; PS00454; FKBP_PPIASE_2; PARTIAL.
 DR PROSITE; PS00059; FKBP_PPIASE_3; PARTIAL.
 KW Isomerase; Rotamase; Repeat; Calmodulin-binding.
 FT NON_TER 1 1
 FT NON_TER 15 15
 FT SEQUENCE 15 AA; 1675 MW; 2B53999722277F3F CRC64;
 SQ

Query Match 52.5%; Score 21; DB 1; Length 15;

Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 GSSFLSP 8
|:|:|
Db 2 GSSFLSP 9

RESULT 10

GP54_BPSPI
ID GP54_BPSPI STANDARD; PRT; 46 AA.
AC 048408;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Putative gene protein 54.
GN 54.
OS Bacteriophage SP01.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC SP01-like viruses.
OX NCBI_TaxID=10685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96327781; PubMed=9657951;
RA Stewart C.R., Gabelightwala I., Hinata K., Krolkoweki K.A.,
RA Needleman D.S., Peng A.S.-Y., Peterman M.A., Tobias A., Wei P.;
RT "Genes and regulatory sites of the 'host-takeover module' in the
terminal redundancy of Bacillus subtilis bacteriophage SP01";
RL Virology 246:329-340(1998).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF031901; AAC29023.1; -
KW Hypothetical protein.

SO SEQUENCE 46 AA; 5288 MW; 0ED5FPA236813246 CRC64;

Query Match 52.5%; Score 21; DB 1; Length 46;
Best Local Similarity 60.0%; Pred. No. 5.7e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 FLSP 8
|:|:|
Db 19 FLSP 23

RESULT 11

OXA_OPHHA
ID OXA_OPHHA STANDARD; PRT; 19 AA.
AC P81383;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUL-1999 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE L-aminic acid oxidase (EC 1.4.3.2) (LAO) (Fragment).
OS Ophiophagus hannah (King cobra) (Naja hannah).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Ophiophagus.
OX NCBI_TaxID=8665;
RN [1]
RP SEQUENCE.
RX TISSUE=Venom;
RX MEDLINE=94361525; PubMed=8080286;
RA Ponudurai G., Chung M.C.M., Tan N.-H.;
RT "Purification and properties of the L-aminic acid oxidase from Malaysian
pit viper (Calloselasma rhodostoma) venom";
RL Arch. Biochem. Biophys. 313:373-378(1994).
RN [2]

RP SEQUENCE OF 1-15.

RC TISSUE=Venom;
RX MEDLINE=97449790; PubMed=9304806;

RA Ahn M.Y., Lee B.M., Kim Y.S.;
RT "Characterization and cytotoxicity of L-aminic acid oxidase from the
venom of king cobra (Ophiophagus hannah).";
RL Int. J. Biochem. Cell Biol. 29:911-919(1997).
CC -1- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: An L-aminic acid + H(2)O + O(2) = a 2-oxo acid
+ NH(3) + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- PTM: GLYCOSYLATED.

CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC STRONG: TO MOUSE Flg-1.

KM Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Venom.
FT CONFLICT 1
FT NON-TER 19 H -> S (IN REF. 2).
SQ SEQUENCE 19 AA; 2298 MW; DD911A5B414F1427 CRC64;

Query Match 50.0%; Score 20; DB 1; Length 19;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 FLSP 8
|:|:|
Db 8 FLSP 13

RESULT 12

TILX_SPIOL
ID TILX_SPIOL STANDARD; PRT; 30 AA.
AC P82537;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thylakoid lumenal 17 kDa protein (P17) (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.

RA Kieselbach T., Bystedt M., Schroeder M.P.;
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
CC -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
KW Chloroplast; Thylakoid.
FT NON-TER 30
SQ SEQUENCE 30 AA; 3335 MW; EBD6462064CB67FF CRC64;

Query Match 50.0%; Score 20; DB 1; Length 30;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GSSFLSP 7
|:|:|
Db 16 GYFLYP 22

RESULT 13

IAPP_BOVIN
ID IAPP_BOVIN STANDARD; PRT; 32 AA.
AC Q28207;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Islet amyloid polypeptide (Amylin) (Fragment).
GN IAPP.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

```

OX NCBI_TaxID=9913;
RN [1]
RA SEQUENCE FROM N.A.
RL Albrandt K., Sierzega M.E., Mull E., Brady E.M.G.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SELECTIVELY INHIBITS INSULIN-STIMULATED GLUCOSE
CC UTILIZATION AND GLYCOGEN DEPOSITION IN MUSCLE, WHILE NOT AFFECTING
CC ADIPOCYTE GLUCOSE METABOLISM.
CC -!- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U62629; AAB05916.1; -
DR InterPro; IPR001693; Calcitonin-like.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; PARTIAL.
KW Hormone; Amyloid.
FT PEPTIDE <1 32 ISLET AMYLOID POLYPEPTIDE.
FT NON TER 32
SQ SEQUENCE 32 AA; 3247 MW; 9A5709394EB44C19 CRC64;

Query Match 50.0%; Score 20; DB 1; Length 32;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFSLP 7
DB 11 ANFLAP 16

RESULT 14
IAPP SHEEP
ID IAPP SHEEP STANDARD; PRT; 32 AA.
AC Q26605;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Islet amyloid polypeptide (Amylin). (Fragment).
GN IAPP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RA SEQUENCE FROM N.A.
RL Albrandt K., Sierzega M.E., Mull E., Brady E.M.G.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SELECTIVELY INHIBITS INSULIN-STIMULATED GLUCOSE
CC UTILIZATION AND GLYCOGEN DEPOSITION IN MUSCLE, WHILE NOT AFFECTING
CC ADIPOCYTE GLUCOSE METABOLISM.
CC -!- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U62629; AAB05916.1; -
DR InterPro; IPR001693; Calcitonin-like.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR SMART; SM00113; CALCITONIN; 1.

```

```

DR PROSITE; PS00258; CALCITONIN; PARTIAL.
KW Hormone; Amyloid.
FT NON TER 1
FT PEPTIDE <1 32 ISLET AMYLOID POLYPEPTIDE.
FT NON TER 32
SQ SEQUENCE 32 AA; 3300 MW; CB5609394EB44C05 CRC64;

Query Match 50.0%; Score 20; DB 1; Length 32;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFSLP 7
DB 11 ANFLAP 16

RESULT 15
ATPO SOLTU
ID ATPO SOLTU STANDARD; PRT; 37 AA.
AC P80504;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase delta chain, mitochondrial [EC 3.6.3.14] (Oligomycin
DE sensitivity conferral protein) (OSCP) (Fragment).
DE Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RA SEQUENCE.
RC TISSUE=Tuber;
RX MEDLINE=97077345; PubMed=8919912;
RA Jansch L., Kruft V., Schmitz U.K., Braun H.P.;
RT "New insights into the composition, molecular mass and stoichiometry
RT of the protein complexes of plant mitochondria.";
RL Plant J. 9:357-368(1996).
CC -!- FUNCTION: THIS PROTEIN SEEMS TO BE PART OF THE STALK THAT LINKS
CC CF(0) TO CF(1). IT EITHER TRANSMITS CONFORMATIONAL CHANGES FROM
CC CF(0) INTO CF(1) OR IS IMPLICATED IN PROTON CONDUCTION.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.
DR InterPro; IPR000711; ATPsynth_OSCP.
DR PROSITE; PS00389; ATPASE_DELTA; PARTIAL.
KW Hydrolase; ATP synthase; CF(1); Hydrogen ion transport;
KW Mitochondrion.
FT NON TER 37
SQ SEQUENCE 37 AA; 4003 MW; 9BFDAB14A298F4AF CRC64;

Query Match 50.0%; Score 20; DB 1; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.4e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSFSLP 7
DB 32 SAFMXP 37

```

Search completed: January 10, 2003, 15:55:47
Job time : 8.27273 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:17 ; Search time 23.6364 Seconds
(without alignments)
69.739 Million cell updates/sec

Title: B
Perfect score: 40
Sequence: 1 gseflspe 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 33835
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	70.0	50	2	Q9ZHW6
2	27	67.5	49	2	Q9KIG0
3	26	65.0	27	12	Q81194
4	26	65.0	27	12	Q81195
5	26	65.0	27	12	Q37295
6	26	65.0	27	12	Q90H84
7	26	65.0	27	12	Q91LB6
8	26	65.0	27	12	Q91LB7
9	26	65.0	27	12	Q91LB8
10	26	65.0	27	12	Q91LB9
11	26	65.0	27	12	Q91LFB
12	26	65.0	27	12	Q91LFC
13	26	65.0	27	12	Q91LFD
14	26	65.0	27	12	Q91LFE
15	26	65.0	27	12	Q91LFF
16	26	65.0	27	12	Q91LFG

17	26	65.0	27	12	Q91LFG	Q91LFG hepatitis c
18	26	65.0	27	12	Q91LFI	Q91LFI hepatitis c
19	26	65.0	27	12	Q91LFL	Q91LFL hepatitis c
20	26	65.0	27	12	Q91LFP	Q91LFP hepatitis c
21	26	65.0	27	12	Q91LFO	Q91LFO hepatitis c
22	26	65.0	27	12	Q91LGI	Q91LGI hepatitis c
23	26	65.0	27	12	Q91LG2	Q91LG2 hepatitis c
24	26	65.0	27	12	Q91LG3	Q91LG3 hepatitis c
25	26	65.0	27	12	Q91LG4	Q91LG4 hepatitis c
26	26	65.0	27	12	Q91LH0	Q91LH0 hepatitis c
27	26	65.0	33	2	Q9R5N9	Q9R5N9 clostridium
28	26	65.0	33	2	Q9R5N7	Q9R5N7 clostridium
29	26	65.0	37	11	Q8VDV2	Q8VDV2 mus musculus
30	26	65.0	40	12	Q91K71	Q91K71 hepatitis c
31	26	65.0	40	12	Q91K21	Q91K21 hepatitis c
32	26	65.0	48	11	Q923S0	Q923S0 mus musculus
33	26	65.0	48	11	Q923R8	Q923R8 rattus norv
34	26	65.0	22	2	Q9R5D9	Q9R5D9 aeromonas h
35	25	62.5	33	12	Q83963	Q83963 avian influ
36	25	62.5	33	12	Q91J38	Q91J38 avian influ
37	25	62.5	36	11	P97644	P97644 rattus norv
38	25	62.5	40	12	Q86793	Q86793 hepatitis c
39	25	62.5	40	12	Q88335	Q88335 hepatitis c
40	25	62.5	45	12	Q68592	Q68592 hepatitis c
41	25	62.5	45	12	Q68593	Q68593 hepatitis c
42	25	62.5	45	12	Q68594	Q68594 hepatitis c
43	25	62.5	45	12	Q92DM3	Q92DM3 hepatitis c
44	24	60.0	20	5	Q9U2M0	Q9U2M0 caenorhabdi
45	24	60.0	37	6	O18822	O18822 sus scrofa

ALIGNMENTS

RESULT 1
Q9ZHW6 PRELIMINARY; PRT; 50 AA.
AC Q9ZHW6; 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major cold shock protein (Fragment).
GN CSPL.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxId=1351;
RN [1] _SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=EF1-1;
RX MEDLINE=98440326; PubMed=9767713;
RA Kim W.S., Khunajakr N., Ren J., Dunn N.W.;
RT "Conservation of the major cold shock protein in lactic acid bacteria".
RL Curr. Microbiol. 37:333-336(1998).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF047608; AAC69999.1; -.
DR HSSP; P32081; ICSP.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR000621; Cold_shock.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Activator; DNA-binding; Transcription regulation.
FT NON_TER 1 50
FT NON_TER 1 50
SQ SEQUENCE 50 AA; 5340 MW; A63FEF118510CF1E CRC64;
Query Match 70.0%; Score 28; DB 2; Length 50;
Best Local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 GSFSLSP 8
      |  |  |  |  |
Db 2 GGFISPE 9

RESULT 2
ID Q9KIG0 PRELIMINARY; PRT; 49 AA.
AC Q9KIG0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase PK-3 (Fragment).
GN PK-3.
OS Streptomyces toyocaensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
ON NCBI_TaxID=55952;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL1509;
RA Neu J.M., Wright G.D.;
RT "Characterization of stopK-1, a novel protein kinase from the
RT glycopeptide antibiotic producer Streptomyces toyocaensis NRRL15009.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF234273; AAF82354.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Kinase.
FT NON_TER 1
FT NON_TER 49
FT NON_TER 49
SQ SEQUENCE 49 AA; 5233 MW; FF73464E8228F824 CRC64;

Query Match 67.5%; Score 27; DB 2; Length 49;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SFLSPE 8
      |  |  |  |  |
Db 44 SYLSPE 49

RESULT 3
ID Q81194 PRELIMINARY; PRT; 27 AA.
AC Q81194;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1090;
RA Kato N., Nakazawa T., Mizutani T., Shimotohno K.;
RT "Susceptibility of human T-lymphotropic virus type I infected cell
RT line MT-2 to hepatitis C virus infection.";
RL Biochem. Biophys. Res. Commun. 206:863-869(1995).
DR EMBL; D43649; BAA07764.1; -.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
DR Kato N., Nakazawa T., Mizutani T., Shimotohno K.;
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
FT NON_TER 27
SQ SEQUENCE 27 AA; 2805 MW; 04170DDCFF3CAA7A CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
      |  |  |  |  |
Db 17 TSFLSP 22

RESULT 4
ID Q81195 PRELIMINARY; PRT; 27 AA.
AC Q81195;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1090;
RA Kato N., Nakazawa T., Mizutani T., Shimotohno K.;
RT "Susceptibility of human T-lymphotropic virus type I infected cell
RT line MT-2 to hepatitis C virus infection.";
RL Biochem. Biophys. Res. Commun. 206:863-869(1995).
DR EMBL; D43650; BAA07765.1; -.
DR Kato N., Nakazawa T., Mizutani T., Shimotohno K.;
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
FT NON_TER 27
SQ SEQUENCE 27 AA; 2777 MW; F570DDCFF3CAA62 CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
      |  |  |  |  |
Db 17 TSFLSP 22

RESULT 5
ID Q37295 PRELIMINARY; PRT; 27 AA.
AC Q37295;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=14;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients.";
RL J. Virol. 72:1640-1646(1998).
DR EMBL; AF018404; AAC03691.1; -.
DR InterPro; IPR002531; HCV NS1.
```

DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM Polypeptide; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2757 MW; 0B97853E5A090ACB CRC64;
Query Match
Best Local Similarity 83.3%; Score 26; DB 12; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 SSFLSP 7
Db 17 ASFLSP 22
RESULT 6
Q9OH84 PRELIMINARY; PRT; 27 AA.
AC Q9OH84; 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Genome polypeptide [contains: envelope glycoprotein E2 (GP68) (GP70)
(NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCB1_Taxid=11103;
GN POL.
RN
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of the hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
RT therapy.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF167044; AAD53677.1; -
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2829 MW; B8370E1ECC1BBD86 CRC64;
Query Match
Best Local Similarity 83.3%; Score 26; DB 12; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 SSFLSP 7
Db 17 TSFLSP 22
RESULT 7
Q9ILB6 PRELIMINARY; PRT; 27 AA.
AC Q9ILB6; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Genome polypeptide [contains: envelope glycoprotein E2 (GP68) (GP70)
(NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCB1_Taxid=11103;
GN POL.
RN
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
RT "Longitudinal analysis of Hepatitis C virus replication and liver

RT fibrosis progression in renal transplant recipients.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207118; AAF75324.1; -
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2672 MW; 3D471891B263D1CA CRC64;
Query Match
Best Local Similarity 83.3%; Score 26; DB 12; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 SSFLSP 7
Db 17 ASFLSP 22
RESULT 8
Q9ILB7 PRELIMINARY; PRT; 27 AA.
AC Q9ILB7; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Genome polypeptide [contains: envelope glycoprotein E2 (GP68) (GP70)
(NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCB1_Taxid=11103;
GN POL.
RN
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
RT "Longitudinal analysis of Hepatitis C virus replication and liver
RT fibrosis progression in renal transplant recipients.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207117; AAF75323.1; -
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2642 MW; 3D471891B273D0CA CRC64;
Query Match
Best Local Similarity 83.3%; Score 26; DB 12; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 SSFLSP 7
Db 17 ASFLSP 22
RESULT 9
Q9ILB8 PRELIMINARY; PRT; 27 AA.
AC Q9ILB8; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Genome polypeptide [contains: envelope glycoprotein E2 (GP68) (GP70)
(NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCB1_Taxid=11103;

```

GN POL
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
Rumeau J.L., Duffaut M., Durand D., Puel J.;
"Longitudinal analysis of Hepatitis C virus replication and liver
fibrosis progression in renal transplant recipients.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF207116; AAF75322.1; -.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2617 MW; 2FF0ED31B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
Db 17 ASFLSP 22

RESULT 10
Q9ILF9
ID Q9ILE9 PRELIMINARY; PRT; 27 AA.
AC Q9ILE9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
DE (NS1)] (Fragment).
GN POL.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
Rumeau J.L., Duffaut M., Durand D., Puel J.;
"Longitudinal analysis of Hepatitis C virus replication and liver
fibrosis progression in renal transplant recipients.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF207115; AAF75321.1; -.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2614 MW; 2FF0E891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
Db 17 ASFLSP 22

RESULT 11
Q9ILF0
ID Q9ILF0 PRELIMINARY; PRT; 27 AA.
AC Q9ILF0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).

```

```

GN POL
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
Rumeau J.L., Duffaut M., Durand D., Puel J.;
"Longitudinal analysis of Hepatitis C virus replication and liver
fibrosis progression in renal transplant recipients.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF207114; AAF75320.1; -.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2614 MW; 2FF0E891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
Db 17 ASFLSP 22

RESULT 12
Q9ILF1
ID Q9ILF1 PRELIMINARY; PRT; 27 AA.
AC Q9ILF1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
DE (NS1)] (Fragment).
GN POL.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
Rumeau J.L., Duffaut M., Durand D., Puel J.;
"Longitudinal analysis of Hepatitis C virus replication and liver
fibrosis progression in renal transplant recipients.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF207113; AAF75319.1; -.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2614 MW; 2FF0E891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSFLSP 7
Db 17 ASFLSP 22

RESULT 13
Q9ILF2
ID Q9ILF2 PRELIMINARY; PRT; 27 AA.
AC Q9ILF2;

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 GN POL.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxId=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
 RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
 RT "Longitudinal analysis of Hepatitis C virus replication and liver
 RT fibrosis progression in renal transplant recipients.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF207112; AAF75316.1; -.
 DR InterPro; IPR002531; HCV_NSI.
 DR Pfam; PF01560; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polypeptide; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2642 MW; 3D471891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSFLSP 7

DB 17 ASFLSP 22

RESULT 14
 O9ILF3 PRELIMINARY; PRT; 27 AA.
 AC O9ILF3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 GN POL.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxId=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
 RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
 RT "Longitudinal analysis of Hepatitis C virus replication and liver
 RT fibrosis progression in renal transplant recipients.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF207111; AAF75317.1; -.
 DR InterPro; IPR002531; HCV_NSI.
 DR Pfam; PF01560; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polypeptide; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2614 MW; 2FF0E891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSFLSP 7

DB 17 ASFLSP 22

RESULT 15
 O9ILF4 PRELIMINARY; PRT; 27 AA.
 AC O9ILF4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 GN POL.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxId=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
 RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
 RT "Longitudinal analysis of Hepatitis C virus replication and liver
 RT fibrosis progression in renal transplant recipients.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF207110; AAF75316.1; -.
 DR InterPro; IPR002531; HCV_NSI.
 DR Pfam; PF01560; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polypeptide; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2642 MW; 3D471891B273D0CA CRC64;

Query Match 65.0%; Score 26; DB 12; Length 27;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSFLSP 7

DB 17 ASFLSP 22

Search completed: January 10, 2003, 15:57:41
 Job time : 24.6364 secs

THIS PAGE BLANK (USPTO)

XX New peptide which antagonises the effect of ghrelin when administered
PT to a mammal
XX
XX Claim 3; Page 6; 9pp; English.
XX The invention relates to a novel peptide which antagonises the effect of
CC ghrelin when administered to a mammal. The peptide is a synthetic
CC analogue of ghrelin. Ghrelin is a 27-28 residue peptide isolated
CC from distinctive cell types in the stomach of rats and humans and has an
CC octanoyl ester attached to a serine residue. Ghrelin is a potent
CC releaser of growth hormone. The peptides are useful for normalising
CC elevated growth hormone levels in mammals such as those suffering
CC from a tumour related to overproduction of growth hormone or acromegaly.
CC The present sequence is a ghrelin antagonising peptide of the
CC invention.
XX
XX Sequence 14 AA;
SQ
Query Match 68.0%; Score 34; DB 23; Length 14;
Best Local Similarity 64.3%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
QY 1 GSSF-----AKLQPR 10
||| |||||
DB 1 GSXFLSPKALQPR 14
RESULT 2
AAB60514
ID AAB60514 standard; peptide; 27 AA.
XX
AC AAB60514;
XX
DT 24-APR-2001 (first entry)
XX
DE Rat des-Gln14-ghrelin, SEQ ID NO:10.
XX
KW Growth hormone secretagogue; GHS; ghrelin;
KW calcium concentration elevation; infant growth disorder;
KW growth hormone deficiency.
XX
OS Rattus norvegicus.
XX
XX WO200107475-A1.
PN
XX 01-FEB-2001.
PD
XX 24-JUL-2000; 2000WO-JP04907.
PF
XX 23-JUL-1999; 99JP-0210002.
PR
XX 29-NOV-1999; 99JP-0338841.
PR
XX 26-APR-2000; 2000JP-0126623.
PR
XX (KANG/) KANGAWA K.
PA
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
PI WPI; 2001-159704/16.
XX
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
XX Claim 3; Page 185; 210pp; Japanese.
XX The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with

CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with
CC no accompanying side effects. The present sequence represents a
CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
XX
XX Sequence 27 AA;
SQ
Query Match 63.0%; Score 31.5; DB 22; Length 27;
Best Local Similarity 37.0%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 0; Indels 17; Gaps 1;
QY 1 GSSF-----AKLQPR 10
||| |||||
DB 1 GSXFLSPKHQAKRKSKPPAKLQPR 27
RESULT 3
AAB60515
ID AAB60515 standard; peptide; 27 AA.
XX
AC AAB60515;
XX
DT 24-APR-2001 (first entry)
XX
DE Human des-Gln14-ghrelin, SEQ ID NO:11.
XX
KW Growth hormone secretagogue; GHS; ghrelin;
KW calcium concentration elevation; infant growth disorder;
KW growth hormone deficiency.
XX
OS Homo sapiens.
XX
XX WO200107475-A1.
PN
XX 01-FEB-2001.
PD
XX 24-JUL-2000; 2000WO-JP04907.
PF
XX 23-JUL-1999; 99JP-0210002.
PR
XX 29-NOV-1999; 99JP-0338841.
PR
XX 26-APR-2000; 2000JP-0126623.
PR
XX (KANG/) KANGAWA K.
PA
XX Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
PI WPI; 2001-159704/16.
XX
XX New peptide compounds which induce growth hormone secretion and
PT elevate cell calcium concentrations, useful in treatment and diagnosis
PT of infant growth disorders -
XX
XX Claim 3; Page 185; 210pp; Japanese.
XX The invention relates to a novel peptide compound or its salt which
CC induces the secretion of growth hormone and/or elevates calcium ion
CC concentration in cells. The peptides are ghrelin homologues and are
CC characterised in that at least one amino acid has been substituted by
CC a modified amino acid and/or a non-amino acid compound. The invention
CC also encompasses the unmodified peptides; the DNA encoding the peptides;
CC vectors and host cells comprising such DNA; a method of producing the
CC peptides comprising recombinant production, optionally followed by
CC chemical modification; an antibody specific for a peptide of the
CC invention; and an assay and kit for detecting the peptides. The peptides
CC of the invention are useful for treating and/or diagnosing diseases
CC caused by a deficiency in growth hormone expression or activity. In
CC particular, they are useful for promoting infant growth due to growth
CC hormone deficiency. The compounds of the invention are safe with

CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 27 AA;

Query Match 63.0%; Score 31.5; DB 22; Length 27;
 Best Local Similarity 37.0%; Pred. No. 21;
 Matches 10; Conservative 0; Mismatches 0; Indels 17; Gaps 1;
 OY 1 GSSF-----AKLQPR 10
 DB 1 GSSFSLPEHQVQRKSKKPPAKLQPR 27

RESULT 4
 AAG64943
 ID AAG64943 standard; peptide; 28 AA.
 AC AAG64943;
 XX
 DT 19-OCT-2001 (first entry)
 XX
 DE Neurone denaturation prevention method related peptide #5.
 XX
 KW Neurone denaturation; neurone death; growth hormone liberation inhibitor;
 KW cerebral infarction; oedema; Alzheimer's disease; Parkinson's disease;
 KW Pick's disease; dementia; amyotrophic lateral sclerosis; cancer;
 KW diabetic neuropathy; neuroprotective; antiinflammatory; nootropic;
 KW cytostatic.
 XX
 OS Unidentified.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 3 /label= OTHER
 FT /note= "modified by O(C=O) (CH2)6CH3"
 FT
 PN WO200147558-A1.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-JP09431.
 XX
 PR 28-DEC-1999; 99JP-0375513.
 XX
 PA (KAKE) KAKEN PHARM CO LTD.
 XX
 PI Murata T, Ohyama T, Amakawa M, Fujita K, Ueo H;
 XX
 DR WPI; 2001-536280/59.
 XX
 PT Agents for treating diseases associated with denaturation or death of
 PT neurons comprise growth hormone liberation inhibitor -
 XX
 PS Disclosure; Page 17; 50pp; Japanese.
 XX
 CC The present invention provides agents for treating or preventing diseases
 CC associated with denaturation or death of neurons, which comprise a
 CC growth hormone liberation inhibitor. These can be used for treating or
 CC preventing diseases associated with denaturation or death of neurons
 CC including those due to cerebral ischaemic disorders such as cerebral
 CC infarction or oedema. Other causes of denaturation or death of neurons
 CC included Alzheimer's disease, Pick's disease, AIDS related dementia,
 CC Parkinson's disease, amyotrophic lateral sclerosis, diabetic neuropathy,
 CC and anticancer treatments. The present sequence is a peptide described in
 CC the exemplification of the invention.
 XX
 SQ Sequence 28 AA;

Query Match 62.0%; Score 31; DB 22; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

OY 1 GSSF-----AKLQPR 10
 DB 1 GSSFSLPEHQKQQRKSKKPPAKLQPR 28

RESULT 5
 AAB60508
 ID AAB60508 standard; peptide; 28 AA.
 AC AAB60508;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Rat ghrelin, SEQ ID NO:2.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 PA (KANG/) KANGAWA K.
 XX
 PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX
 DR WPI; 2001-159704/16.
 XX
 PT New peptide compounds which induce growth hormone secretion and
 PT elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 XX
 PS Claim 2; Page 180; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 28 AA;

Query Match 62.0%; Score 31; DB 22; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

OY 1 GSSF-----AKLQPR 10
 DB 1 GSSFSLPEHQKQQRKSKKPPAKLQPR 28

RESULT 6
 AAB60509

ID AAB60509 standard; peptide; 28 AA.
 XX
 AC AAB60509;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human ghrelin, SEQ ID NO:3.
 XX
 KW Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 PA (KANG/) KANGAWA K.
 XX
 PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX
 DR WPI; 2001-159704/16.
 XX
 CC New peptide compounds which induce growth hormone secretion and
 CC elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 PT
 PS Claim 3; Page 181; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 28 AA;
 Query Match 62.0%; Score 31; DB 22; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
 QY 1 GSSF-----AKLQPR 10
 |||||
 Db 1 GSSFSLSPHQVQQRKESKPKLQPR 28
 RESULT 7
 AAB60530
 ID AAB60530 standard; peptide; 28 AA.
 XX
 AC AAB60530;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Dog ghrelin-like GH secretagogue peptide, SEQ ID NO:31.

XX
 KW Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.
 XX
 OS Canis familiaris.
 XX
 PN WO200107475-A1.
 XX
 PD 01-FEB-2001.
 XX
 PF 24-JUL-2000; 2000WO-JP04907.
 XX
 PR 23-JUL-1999; 99JP-0210002.
 PR 29-NOV-1999; 99JP-0338841.
 PR 26-APR-2000; 2000JP-0126623.
 XX
 PA (KANG/) KANGAWA K.
 XX
 PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;
 XX
 DR WPI; 2001-159704/16.
 XX
 CC New peptide compounds which induce growth hormone secretion and
 CC elevate cell calcium concentrations, useful in treatment and diagnosis
 PT of infant growth disorders -
 PT
 PS Claim 4; Page 197; 210pp; Japanese.
 XX
 CC The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by
 CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.
 XX
 SQ Sequence 28 AA;
 Query Match 62.0%; Score 31; DB 22; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
 QY 1 GSSF-----AKLQPR 10
 |||||
 Db 1 GSSFSLSPHQVQQRKESKPKLQPR 28
 RESULT 8
 AAE19032
 ID AAE19032 standard; peptide; 28 AA.
 XX
 AC AAE19032;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human ghrelin peptide analogue, compound 6.
 XX
 KW Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
 KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
 KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
 KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
 KW cardiovascular disorder; gall stone; osteoarthritis; cancer; cytostatic;
 KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;

KW cardiant; litholytic; hepatotropic.
 XX
 OS Homo sapiens.
 XX
 PN WO200192292-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 25-MAY-2001; 2001WO-US17026.
 XX
 PR 30-MAY-2000; 2000US-207920P.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Bednarek M;
 XX
 DR WPI; 2002-195531/25.
 XX
 PT Truncated ghrelin analogs active at growth-hormone secretagogue
 PT receptor useful for diagnosing or treating diseases such as anorexia,
 PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
 PT osteoarthritis
 XX
 PS Example 4; Page 34; 37pp; English.
 XX
 CC The present invention relates to a truncated ghrelin analogue or their
 CC salt, active at growth-hormone secretagogue (GHS) receptor. Ghrelin
 CC analogue is useful for screening a compound capable of binding to GHS
 CC receptor and for stimulating growth hormone secretion. Ghrelin agonist
 CC is utilized for treating a growth hormone deficient state, increasing
 CC muscle mass and bone density, treating sexual dysfunction in males or
 CC females, facilitating a weight gain, maintenance of weight, maintenance
 CC of physical functioning, recovery of physical function, and/or appetite
 CC increase, or appetite increase is particularly useful for a patient
 CC having a disease or disorder, or under going a treatment, accompanied by
 CC weight loss such as anorexia, bulimia, cancer cachexia, acquired
 CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
 CC elderly and examples of treatments accompanied by weight loss include
 CC chemotherapy, radiation therapy, temporary or permanent immobilisation
 CC and dialysis; and ghrelin antagonist is utilised to facilitate weight
 CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
 CC complications of diabetes including retinopathy, and/or cardiovascular
 CC disorders, where excessive weight is a contributing factor to different
 CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
 CC disease, gall stones, osteoarthritis and certain forms of cancers, and
 CC bringing about a weight loss can be used for e.g. to reduce the
 CC likelihood of such diseases and for treating such diseases. Ghrelin
 CC analogue induces growth hormone release from primary-culture pituitary
 CC cells in a dose-dependent manner without stimulating the release of other
 CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
 CC synthesised easily and has increased solubility in physiological buffers.
 CC The present sequence is human ghrelin peptide analogue.
 XX
 SQ Sequence 28 AA;
 XX
 Query Match Score 31; DB 23; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
 QY 1 GSSF-----AKLQPR 10
 DB 1 GGSFSLPEHQRVQORKESSKPPAKLQPR 28
 XX
 RESULT 9
 AAE19040
 ID AAE19040 standard; peptide; 28 AA.
 XX
 AC AAE19040;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human ghrelin peptide analogue, compound 16.

XX
 KW Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
 KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
 KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
 KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
 KW cardiovascular disorder; gall stone; osteoarthritis; cancer; cytostatic;
 KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;
 KW cardiant; litholytic; hepatotropic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 3
 FT /note= "Ser(CO-(CH2)5-CH3)"
 XX
 PN WO200192292-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 25-MAY-2001; 2001WO-US17026.
 XX
 PR 30-MAY-2000; 2000US-207920P.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Bednarek M;
 XX
 DR WPI; 2002-195531/25.
 XX
 PT Truncated ghrelin analogs active at growth-hormone secretagogue
 PT receptor useful for diagnosing or treating diseases such as anorexia,
 PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
 PT osteoarthritis
 XX
 PS Example 4; Page 23; 37pp; English.
 XX
 CC The present invention relates to a truncated ghrelin analogue or their
 CC salt, active at growth-hormone secretagogue (GHS) receptor. Ghrelin
 CC analogue is useful for screening a compound capable of binding to GHS
 CC receptor and for stimulating growth hormone secretion. Ghrelin agonist
 CC is utilized for treating a growth hormone deficient state, increasing
 CC muscle mass and bone density, treating sexual dysfunction in males or
 CC females, facilitating a weight gain, maintenance of weight, maintenance
 CC of physical functioning, recovery of physical function, and/or appetite
 CC increase, or appetite increase is particularly useful for a patient
 CC having a disease or disorder, or under going a treatment, accompanied by
 CC weight loss such as anorexia, bulimia, cancer cachexia, acquired
 CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
 CC elderly and examples of treatments accompanied by weight loss include
 CC chemotherapy, radiation therapy, temporary or permanent immobilisation
 CC and dialysis; and ghrelin antagonist is utilised to facilitate weight
 CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
 CC complications of diabetes including retinopathy, and/or cardiovascular
 CC disorders, where excessive weight is a contributing factor to different
 CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
 CC disease, gall stones, osteoarthritis and certain forms of cancers, and
 CC bringing about a weight loss can be used for e.g. to reduce the
 CC likelihood of such diseases and for treating such diseases. Ghrelin
 CC analogue induces growth hormone release from primary-culture pituitary
 CC cells in a dose-dependent manner without stimulating the release of other
 CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
 CC synthesised easily and has increased solubility in physiological buffers.
 CC The present sequence is human ghrelin peptide analogue.
 XX
 SQ Sequence 28 AA;
 XX
 Query Match Score 31; DB 23; Length 28;
 Best Local Similarity 35.7%; Pred. No. 28;
 Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
 QY 1 GSSF-----AKLQPR 10
 DB 1 GGSFSLPEHQRVQORKESSKPPAKLQPR 28
 XX

RESULT 10
AAE19041
ID AAE19041 standard; peptide; 28 AA.
XX
AC AAE19041;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human ghrelin peptide analogue, compound 17.
XX
KW Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
KW cardiovascular disorder; gall stone; osteoarthritis; cancer; cytostatic;
KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;
KW cardiant; litholytic; hepatotropic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Ser (CO-(CH2)6-CH3)"
XX
PN WO200192292-A2.
XX
PD 06-DEC-2001.
XX
PF 25-MAY-2001; 2001WO-US17026.
XX
PR 30-MAY-2000; 2000US-207920P.
XX
PA (MERI) MERCK & CO INC.
XX
PI Bednarek M;
XX
DR WPI; 2002-195531/25.
XX
PT Truncated ghrelin analogs active at growth-hormone secretagogue
PT receptor useful for diagnosing or treating diseases such as anorexia,
PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
PT osteoarthritis
XX
PS Example 4; Page 23; 37pp; English.
XX
CC The present invention relates to a truncated ghrelin analogue or their
CC salt, active at growth-hormone secretagogue (GHS) receptor. Ghrelin
CC analogue is useful for screening a compound capable of binding to GHS
CC receptor and for stimulating growth hormone secretion. Ghrelin agonist
CC is utilised for treating a growth hormone deficient state, increasing
CC muscle mass and bone density, treating sexual dysfunction in males or
CC females, facilitating a weight gain, maintenance of weight, maintenance
CC of physical functioning, recovery of physical function, and/or appetite
CC increase, or appetite increase is particularly useful for a patient
CC having a disease or disorder, or under going a treatment, accompanied by
CC eight loss such as anorexia, bulimia, cancer cachexia, acquired
CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
CC elderly and examples of treatments accompanied by weight loss include
CC chemotherapy, radiation therapy, temporary or permanent immobilisation
CC and dialysis; and ghrelin antagonist is utilised to facilitate weight
CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
CC complications of diabetes including retinopathy, and/or cardiovascular
CC disorders, where excessive weight is a contributing factor to different
CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
CC diseases, gall stones, osteoarthritis and certain forms of cancers, and
CC bringing about a weight loss can be used for e.g. to reduce the
CC likelihood of such diseases and for treating such diseases. Ghrelin
CC analogue induces growth hormone release from primary-culture pituitary
CC cells in a dose-dependent manner without stimulating the release of other
CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
CC synthesised easily and has increased solubility in physiological buffers.

CC The present sequence is human ghrelin peptide analogue.

XX SQ Sequence 28 AA;

Query Match 62.0%; Score 31; DB 23; Length 28;
Best Local Similarity 35.7%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY 1 GSSF-----AKLQPR 10
|||
Db 1 GSSFLSPFHQRVQORKEKKPPAKLQPR 28

RESULT 11

AAAY40027

ID AAY40027 standard; Peptide; 40 AA.

XX

AC AAY40027;

XX

DT 18-NOV-1999 (first entry)

XX

DE Peptide sequence derived from a human secreted protein.

XX

KW Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
KW neurodegenerative disorder; developmental abnormality; blood disorder;
KW immune system disease; autoimmune disease; leukemia; inflammation;
KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
KW connective tissue disorder; transplant rejection; sepsis; acne;
KW psoriasis; cardiovascular disorder; reproductive disorder;
KW food additive; food preservative; storage capability.

XX Homo sapiens.

XX

PN WO9943693-A1.

XX

PD 02-SEP-1999.

XX

PF 24-FEB-1999; 99WO-US03939.

XX

PR 26-FEB-1998; 98US-0076051.

XX

PR 26-FEB-1998; 98US-0076052.

XX

PR 26-FEB-1998; 98US-0076053.

XX

PR 26-FEB-1998; 98US-0076054.

XX

PR 26-FEB-1998; 98US-0076057.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PI Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;

XX Duan RD;

XX

PI WPI; 1999-550857/46.

XX

DR New human genes and the secreted polypeptides they encode, useful for

XX diagnosis and treatment of e.g. cancers, neurological disorders, immune

XX diseases, inflammation or blood disorders

XX

PS Disclosure; Page 27; 246pp; English.

XX

XX AAY40001-92 are derived from human secreted proteins. The

XX polynucleotides and their corresponding secreted polypeptides are useful

XX for preventing, treating or ameliorating medical conditions, e.g. by

XX protein or gene therapy. Pathological conditions can also be diagnosed by

XX determining the amount of the new polypeptides in a sample or by

XX determining the presence of mutations in the polynucleotide. Specific

XX uses include developing products for the diagnosis or treatment of

XX cancer, tumors, neurodegenerative disorders, developmental abnormalities

XX and fetal deficiencies, blood disorders, sepsis, diseases of the immune

XX system, autoimmune diseases, inflammation, allergies, Alzheimer's and

XX cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,

XX infections, AIDS, connective tissue disorders, transplant rejection,

XX diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,

XX and reproductive disorders. The polypeptides or polynucleotides can

CC also be used as food additives or preservatives, such as to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components.

CC Sequence 40 AA;

Query Match 62.0%; Score 31; DB 20; Length 40;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSGFAKLP 9
 |||||
 Db 15 GSGFAKLP 23

RESULT 12
 AAU99715
 ID AAU99715 standard; Peptide; 12 AA.

AC AAU99715;

DT 24-SEP-2002 (first entry)

DE Human Ghrelin (C-terminus) peptide sequence.

XX Human; angiotensin converting enzyme-2; ACE-2; body weight disorder;
 KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;
 KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;
 KW familial partial lipodystrophy; hypercholesterolaemia; hyperlipidaemia;
 KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;
 KW neurodegenerative disorder; peptide hormone; cytokine processing;
 KW myocardial infarction; cardiomyopathy; inflammatory bowel disease;
 KW systemic inflammation response syndrome; polycytemia; pain; stroke;
 KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;
 KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;
 KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;
 KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;
 KW adult respiratory distress syndrome; wound healing; appetite;
 KW body mass index; Ghrelin.

XX Homo sapiens.

PN WO20023997-A2.

XX 23-MAY-2002.

PF 31-OCT-2001; 2001WO-US45703.

PR 01-NOV-2000; 2000US-0704216.

PR 29-MAY-2001; 2001US-0870382.

PR 19-OCT-2001; 2001US-371741P.

(MILL-) MILLENNIUM PHARM INC.

PI Action SL, Ocalin TD, Gould AE, Dales NA, Guan B, Brown JA;

PI Patene M, Kadambi VJ, Solomon M, Stricker-Krongrad A;

DR WPI; 2002-547572/58.

PT Treating body weight disorder and increasing muscle mass comprises

PT administering angiotensin converting enzyme-2 modulating compound -

PS Example 18; Page 221; 395pp; English.

XX The present invention describes a new method of creating a body weight
 CC disorder, increasing muscle mass and decreasing body fat by
 CC administration of angiotensin converting enzyme (ACE)-2 modulating
 CC compound. The invention can be used for treating body weight disorders,
 CC particularly obesity of at least grade 1, diabetes, atherosclerosis and
 CC a state associated with lipid metabolism. The method is used for treating
 CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,
 CC generalised partial lipodystrophy, familial partial lipodystrophy,

CC hypercholesterolaemia, hyperlipidaemia, an aberrant metabolic rate,
 CC congestive heart failure, chronic heart failure, left ventricular
 CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease and Huntington's disease),
 CC diseases associated with peptide hormones or cytokine processing,
 CC myocardial infarction, cardiomyopathy, systemic inflammation response
 CC syndrome, sepsis, polycytemia, inflammatory bowel disease, acute and
 CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis
 CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema
 CC following focal injury, diffuse axonal injury, stroke, reperfusion
 CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic
 CC disorders including asthma, adult respiratory distress syndrome, wound
 CC healing and scar formation. The invention decreases the appetite,
 CC increases muscle mass and decreases body fat of subject having body mass
 CC index of greater than 23 (preferably 24.9)kg/m². The present amino
 CC acid sequence represents the human Ghrelin (C-terminus) peptide that was
 CC used in the invention for hydrolysis of biologically active peptides
 CC by soluble human ACE-2.

Sequence 12 AA;

Query Match 60.0%; Score 30; DB 23; Length 12;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKLQPR 10
 |||||
 Db 7 AKLQPR 12

RESULT 13
 AAB60560
 ID AAB60560 standard; peptide; 28 AA.

XX AAB60560;

DT 25-APR-2001 (first entry)

DE Rat ghrelin-derived growth hormone secretagogue (GHS) peptide.

XX Growth hormone secretagogue; GHS; ghrelin;
 KW calcium concentration elevation; infant growth disorder;
 KW growth hormone deficiency.

OS Rattus norvegicus.

OS Synthetic.

PN WO200107475-A1.

PD 01-FEB-2001.

PF 24-JUL-2000; 2000WO-JP04907.

PR 23-JUL-1999; 99JP-0210002.

PR 29-NOV-1999; 99JP-0338841.

PR 26-APR-2000; 2000JP-0126623.

(KANG/) KANGAWA K.

PI Kangawa K, Kojima M, Hosoda H, Matsuo H, Minamitake Y;

DR WPI; 2001-159704/16.

PT New peptide compounds which induce growth hormone secretion and

PT elevate cell calcium concentrations, useful in treatment and diagnosis

PS Example 3; Page 84; 210pp; Japanese.

XX The invention relates to a novel peptide compound or its salt which
 CC induces the secretion of growth hormone and/or elevates calcium ion
 CC concentration in cells. The peptides are ghrelin homologues and are
 CC characterised in that at least one amino acid has been substituted by

CC a modified amino acid and/or a non-amino acid compound. The invention
 CC also encompasses the unmodified peptides; the DNA encoding the peptides;
 CC vectors and host cells comprising such DNA; a method of producing the
 CC peptides comprising recombinant production, optionally followed by
 CC chemical modification; an antibody specific for a peptide of the
 CC invention; and an assay and kit for detecting the peptides. The peptides
 CC of the invention are useful for treating and/or diagnosing diseases
 CC caused by a deficiency in growth hormone expression or activity. In
 CC particular, they are useful for promoting infant growth due to growth
 CC hormone deficiency. The compounds of the invention are safe with
 CC no accompanying side effects. The present sequence represents a
 CC ghrelin-type growth hormone secretagogue (GHS) of the invention.

XX Sequence 28 AA;

Query Match 60.0%; Score 30; DB 22; Length 28;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKLQPR 10
 |||||
 Db 23 AKLQPR 28

RESULT 14

AAE19021
 ID AAE19021 standard; peptide; 28 AA.

AC AAE19021;

DT 21-MAY-2002 (first entry)

DE Human ghrelin peptide.

XX Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
 KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
 KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
 KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
 KW cardiovascular disorder; gall stone; osteoarthritis; cancer; cytostatic;
 KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;
 KW cardiant; litholytic; hepatotropic.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 3
 FT /note= "Ser(CO-(CH2)6-CH3)"

XX WO200192292-A2.

XX 06-DEC-2001.

XX 25-MAY-2001; 2001WO-US17026.

XX 30-MAY-2000; 2000US-207920P.

XX (MERI) MERCK & CO INC.

XX Bednarek M;

XX WPI; 2002-195531/25.

XX Truncated ghrelin analogs active at growth-hormone secretagogue

PT receptor useful for diagnosing or treating diseases such as anorexia,
 PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
 PT osteoarthritis -

XX Disclosure; Page 3; 37pp; English.

XX The present invention relates to a truncated ghrelin analogue or their
 CC salt; active at growth-hormone secretagogue (GHS) receptor. Ghrelin
 CC analogue is useful for screening a compound capable of binding to GHS
 CC receptor and for stimulating growth hormone secretion. Ghrelin agonist

CC is utilised for treating a growth hormone deficient state, increasing
 CC muscle mass and bone density, treating sexual dysfunction in males or
 CC females, facilitating a weight gain, maintenance of weight, maintenance
 CC of physical functioning, recovery of physical function, and/or appetite
 CC increase, or appetite increase is particularly useful for a patient
 CC having a disease or disorder, or under going a treatment, accompanied by
 CC eight loss such as anorexia, bulimia, cancer cachexia, acquired
 CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
 CC elderly and examples of treatments accompanied by weight loss include
 CC chemotherapy, radiation therapy, temporary or permanent immobilisation
 CC and dialysis; and ghrelin antagonist is utilised to facilitate weight
 CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
 CC complications of diabetes including retinopathy, and/or cardiovascular
 CC disorders, where excessive weight is a contributing factor to different
 CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
 CC disease, gall stones, osteoarthritis and certain forms of cancers, and
 CC bringing about a weight loss can be used for e.g. to reduce the
 CC likelihood of such diseases and for treating such diseases. Ghrelin
 CC analogue induces growth hormone release from primary-culture pituitary
 CC cells in a dose-dependent manner without stimulating the release of other
 CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
 CC synthesised easily and has increased solubility in physiological buffers.
 CC The present sequence is human ghrelin peptide.

XX Sequence 28 AA;

Query Match 60.0%; Score 30; DB 23; Length 28;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKLQPR 10

|||||

Db 23 AKLQPR 28

RESULT 15

AAE19027
 ID AAE19027 standard; peptide; 28 AA.

XX AAE19027;

XX 21-MAY-2002 (first entry)

XX Human ghrelin peptide analogue, compound 1.

XX Human; ghrelin analogue; growth-hormone secretagogue; GHS receptor; AIDS;
 KW acquired immune deficiency syndrome; weight gain; chemotherapy; dialysis;
 KW growth hormone; muscle mass; bone density; sexual dysfunction; anorexia;
 KW wasting; radiation therapy; obesity; diabetes; retinopathy; hypertension;
 KW cardiovascular disorder; gall stone; osteoarthritis; cancer; cytostatic;
 KW metabolic; immunomodulator; anti-HIV; anorectic; ophthalmological;
 KW cardiant; litholytic; hepatotropic.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 3
 FT /note= "Ser(CO-CH=CH-CH=CH-CH3)"

XX WO200192292-A2.

XX 06-DEC-2001.

XX 25-MAY-2001; 2001WO-US17026.

XX 30-MAY-2000; 2000US-207920P.

XX (MERI) MERCK & CO INC.

XX Bednarek M;

XX WPI; 2002-195531/25.

PT Truncated ghrelin analogs active at growth-hormone secretagogue
 PT receptor useful for diagnosing or treating diseases such as anorexia,
 PT bulimia, cancer, obesity, diabetes mellitus, hypertension,
 PT osteoarthritis -

PS Example 4; Page 33; 37pp; English.

XX
 CC The present invention relates to a truncated ghrelin analogue or their
 CC salt, active at growth-hormone secretagogue (GHS) receptor. Ghrelin
 CC analogue is useful for screening a compound capable of binding to GHS
 CC receptor and for stimulating growth hormone secretion. Ghrelin agonist
 CC is utilized for treating a growth hormone deficient state, increasing
 CC muscle mass and bone density, treating sexual dysfunction in males or
 CC females, facilitating a weight gain, maintenance of weight, maintenance
 CC of physical functioning, recovery of physical function, and/or appetite
 CC increase, or appetite increase is particularly useful for a patient
 CC having a disease or disorder, or under going a treatment, accompanied by
 CC eight loss, such as anorexia, bulimia, cancer cachexia, acquired
 CC immune deficiency syndrome (AIDS), wasting, cachexia and wasting in frail
 CC elderly and examples of treatments accompanied by weight loss include
 CC chemotherapy, radiation therapy, temporary or permanent immobilization
 CC and dialysis; and ghrelin antagonist is utilized to facilitate weight
 CC loss, appetite decrease, weight maintenance, treat obesity, diabetes and
 CC complications of diabetes including retinopathy, and/or cardiovascular
 CC disorders, where excessive weight is a contributing factor to different
 CC diseases including hypertension, diabetes, dyslipidemias, cardiovascular
 CC disease, gall stones, osteoarthritis and certain forms of cancers, and
 CC bringing about a weight loss can be used for e.g. to reduce the
 CC likelihood of such diseases and for treating such diseases. Ghrelin
 CC analogue induces growth hormone release from primary-culture pituitary
 CC cells in a dose-dependent manner without stimulating the release of other
 CC pituitary hormones. Unlike longer length ghrelin, ghrelin analogue can be
 CC synthesised easily and has increased solubility in physiological buffers.
 CC The present sequence is human ghrelin peptide analogue.

XX
 SQ Sequence 28 AA;

Query Match 60.0%; Score 30; DB 23; Length 28;

Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKLQPR 10
 |||||
 |||||
 Db 23 AKLQPR 28

Search completed: January 10, 2003, 15:59:14
 Job time : 40.0909 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 10, 2003, 15:55:16 ; Search time 15.9091 Seconds
(without alignments)
60.427 Million cell updates/sec

Title: C
Perfect score: 50
Sequence: 1 gsefaklqpr 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 11827

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	64.0	43	2	S41388 protein 3a - human
2	28	56.0	16	2	IS1879 cyathionine beta
3	28	56.0	19	2	T02624 hypothetical prote
4	27	54.0	27	1	RHPGA gastrin-releasing
5	27	54.0	27	1	RHPGA gastrin-releasing
6	27	54.0	37	2	D83199 hypothetical prote
7	26	52.0	39	2	G89467 protein R09H3.2 [l
8	25	50.0	11	2	S66606 quinioline 2-oxid
9	25	50.0	25	2	S06263 gastrin-releasing
10	25	50.0	29	2	B61613 ceratotoxin B - Me
11	25	50.0	43	2	A10982 hypothetical prote
12	25	50.0	49	2	B97874 degenerate transp
13	24	48.0	30	2	S08565 ribulose-bisphosph
14	24	48.0	39	2	A47752 RNA recognition mo
15	24	48.0	48	2	P00617 DNA-directed DNA p
16	24	48.0	50	2	B82279 hypothetical prote
17	23	46.0	19	2	A34467 36K microfilbril-as
18	23	46.0	27	2	P00844 DNA-binding protei
19	23	46.0	29	2	S02578 H+-transporting tw
20	23	46.0	30	2	D81561 hypothetical prote
21	23	46.0	36	2	E41080 rbpl protein - Rho
22	23	46.0	37	2	C70187 hypothetical prote
23	23	46.0	39	2	B81912 hypothetical prote
24	23	46.0	40	2	A99799 hypothetical prote
25	23	46.0	43	2	I45824 aldolase C - New Z
26	23	46.0	44	2	A60329 antigen PV9 - Plas
27	23	46.0	45	2	A81796 hypothetical prote
28	23	46.0	50	2	E70176 hypothetical prote
29	22	44.0	13	2	PH1316 Ig heavy chain DJ

30	22	44.0	22	2	C60691 phycobillosome 99K
31	22	44.0	31	2	S31176 microtubule-associ
32	22	44.0	35	2	B83824 hypothetical prote
33	22	44.0	39	2	AF2779 hypothetical prote
34	22	44.0	40	2	S19539 triacylglycerol 1i
35	22	44.0	43	2	I46150 aldolase C - dog (
36	22	44.0	44	2	hypothetical prote
37	22	44.0	47	1	W0BP57 gene 0.5 protein -
38	21	42.0	7	2	A58718 carnosin U149 - Ca
39	21	42.0	13	2	S47358 T-cell antigen rec
40	21	42.0	13	2	S78766 ribosomal protein
41	21	42.0	16	2	I40065 shikimate 5-dehydr
42	21	42.0	26	2	S33869 ribosomal protei
43	21	42.0	27	2	A38123 probable maud prot
44	21	42.0	29	2	I37301 MHC class II histo
45	21	42.0	29	2	I37303 HLA-DR beta - huma

ALIGNMENTS

RESULT 1

S41388
protein 3a - human adenovirus 3 (fragment)
C/Species: Mastadenovirus h3 (human adenovirus 3)
C/Date: 06-Jan-1995 #sequence_revision 30-Jan-1998 #text_change 26-Aug-1999
C/Accession: S41388
R/Cuzange, A.; Chroboczek, J.; Jacrot, B.
submitted to the EMBL Data Library, January 1994
A/Description: The penton base of human adenovirus type 3 has the RGD motif.
A/Reference number: S41388
A/Accession: S41388
A/Molecule type: DNA
A/Residues: 1-43 <CUZ>

A/Cross-references: EMBL:Z29487, NID:g444048, PIDN:CAA82621.1, PID:g444049
A/Experimental source: serotype 3
A/Superfamily: adenovirus pentononal hexon-associated protein

Query Match

Best Local Similarity 64.0%; Score 32; DB 2; Length 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 10
Db 29 GNPFAHLRPR 38

RESULT 2

IS1879
cyathionine beta-synthase - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C/Accession: IS1879
R/Sebastio, G.; Sperandio, M.P.; Panico, M.; de Franchis, R.; Kraus, J.P.; Andria, G.
Am. J. Hum. Genet. 56, 1324-1333, 1995
A/Title: The molecular basis of homocystinuria due to cyathionine beta-synthase deficit
A/Reference number: IS1879; PMID:95282779; PMID:7762555
A/Accession: IS1879
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <RSS>
A/Cross-references: GB:S78267, NID:g999349, PIDN:AA834404.1, PID:g999350

Query Match

Best Local Similarity 56.0%; Score 28; DB 2; Length 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 9
Db 8 GGAFKLQPR 16

RESULT 3

T02624
hypothetical protein At2g25990 [imported] - Arabidopsis thaliana
N;Alternate names: hypotheical protein T19L18.20
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Feb-2001
C;Accession: T02624; B84655
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, August 1998
A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.
A;Reference number: Z14681
A;Accession: T02624
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-19 <ROU>
A;Cross-references: EMBL:AC004747; NID:G3413696; PID:G3413713
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84655
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-19 <STO>
A;Cross-references: GB:AE002093; NID:G3413713; PIDN:AA31236.1; GSPDB:GN00139
C;Genetics:
A;Map position: 2
A;Gene: At2g25990; T19L18.20

```
Query Match          56.0%; Score 28; DB 2; Length 19;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

Qy 1 GSSFAKLQPR 10
Dp 3 GSOREKLKPR 12

RESULT 4
RHPGA
gastrin-releasing peptide - pig
N; Contains: neuromedin C
C; Species: Sus scrofa domestica (domestic pig)
C; Date: 31-Mar-1981 #sequence, revision 31-Mar-1981 #text_change 20-Mar-1998
C; Accession: A01562; A32738; C60301
R; McDonald, T. J.; Jorvall, H.; Nilsson, G.; Vagne, M.; Gbatei, M.; Bloom, S. R.; Mutt, V.
Biochem. Biophys. Res. Commun. 90, 227-233, 1979
A; Title: Characterization of a gastrin releasing peptide from porcine non-antral gastric
A; Reference number: A01562, MUID:80042433; PMID:456973
A; Accession: A01562
A; Molecule type: protease
A; Residues: 1-27 <MCD>
R; Minamino, N.; Kangawa, K.; Matsuo, H.
Biochem. Biophys. Res. Commun. 119, 14-20, 1984
A; Title: Neuromedin C: a bombesin-like peptide identified in porcine spinal cord.
A; Reference number: A32738; MUID:84153890; PMID:6546686

Accession: A32738
Molecule type: protein
Residues: 18-27 <MIN>
Superfamily: gastrin-releasing peptide
Keywords: amidated carboxyl end; hormone; neuropeptide; spinal cord; stomach
1-27/Product: gastrin-releasing peptide #status experimental <MAT>
18-27/Product: neuromedin C #status experimental <NEU>
27/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 54.0%; Score 27; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GSSFAKLQPR 10
| : | : |

Db 8 GTVLAKMYPR 17

RESULT 5
RHGPGA

gastrin-releasing peptide - guinea pig
 N/Contains: neuromedin C
 C/Species: Cavia porcellus (guinea pig)
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 20-Mar-1998
 C/Accession: A60206

R;Shaw, C.; Thim, L.; Conlon, J.M.
J. Neurochem. 49, 1348-1354, 1987
A;Title: Primary structure and tissue distribution of guinea pig gastrin-releasing pept
A;Reference number: A60206; MUID:88034997; PMID:3668528
A;Accession: A60206
A;Molecule type: protein
A;Residues: 1-27 <SHA>
A;Note: the carboxyl-terminal residue was not determined directly, but an alpha-amidate
ion times
C;Superfamily: gastrin-releasing peptide
C;Keywords: amidated carboxyl end; brain; hormone; intestine; neuropeptide; stomach
F;18-27/Product: neuromedin C #status predicted <NEU>
F;27/Modified site: amidated carboxyl end (Met) #status predicted

QY 1 GSSFAKLQPR 10
| : | : | : | :
Db 8 GTVLAKMYPR 17

RESULT 6
D83199
hypothetical protein PA3577 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83199
F:Stover, C.K.; Yam, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickes-
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lard-
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportu-
nistic pathogen. PMID:10984043
A:Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: D83199
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <STO>
A;Cross-references: GB:AE004778; GB:AE004091; NID:g9949722; PID:AAG06965.1; GS:
A;Experimental source: strain PAQ1
C;Genetics:
A;Gene: PA3577

Query Match	54.0%	Score 27;	DB 2;	Length 37;
Best Local Similarity	62.5%	Pred. No. 60;		
Matches 5;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 GSSFAKLQ 8
Db 7 GESFGRLQ 14

```

RESULT 7
G89467
protein R09H3.2 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C:Accession: G89467
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: A platform for investi
A:Reference number: A75000; MUID:99069613; PMID:9851916

```

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89467
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <STO>
A:Cross-references: GB:chr_X; PIDN:AAB00612.1; PID:g1326316; GSPDB:GN00028; CESP:R09H3.2
C:Genetics:
A:Gene: R09H3.2
A:Map position: X

Query Match 52.0%; Score 26; DB 2; Length 39;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AKQPR 10
Db 2 AKQPR 7

RESULT 8

S66606
quinoline 2-oxidoeductase alpha chain - Comamonas testosteroni (fragment)
C:Species: Comamonas testosteroni
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S66606
R:Schach, S.; Tshisuka, B.; Fetzner, S.; Lingens, F.
Burr, J. Biochem 232, 536-544, 1995
A:Title: Quinolone 2-oxidoeductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase from
A:Reference number: S66606; MUID:96035869; PMID:7556204
A:Accession: S66606
A:Molecule type: protein
A:Residues: 1-11 <SCH>
A:Experimental source: strain 63

Query Match 50.0%; Score 25; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSFAKLQPR 10
Db 3 SDVAKLQPR 11

RESULT 9

S06263
gastrin-releasing peptide - smaller spotted catshark (tentative sequence) (fragment)
C:Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
C:Date: 31-Mar-1990 #sequence_revision 30-Jan-1998 #text_change 31-Mar-2000
C:Accession: S06263
R:Conlon, J.M.; Henderson, I.W.; Thim, L.
Gen. Comp. Endocrinol. 68, 415-420, 1987
A:Title: Gastrin-releasing peptide from the intestine of the elasmobranch fish, Scyliorhinus
A:Reference number: S06263; MUID:88137922; PMID:3436516
A:Accession: S06263
A:Molecule type: protein
A:Residues: 1-25 <CON>
A>Note: the sequence from the summary is inconsistent with that from table 1 and table 3
C:Superfamily: gastrin-releasing peptide
C:Keywords: neuropeptide

Query Match 50.0%; Score 25; DB 2; Length 25;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SPFAKLQPR 10
Db 8 SPFAKLQPR 15

RESULT 10

B61613
ceratotoxin B - Mediterranean fruit fly

C:Species: Ceratitis capitata (Mediterranean fruit fly)
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 17-Mar-1999
C:Accession: B61613
R:Marchini, D.; Giordano, P.C.; Amons, R.; Benini, L.F.; Dallai, R.
Insect Biochem. Mol. Biol. 23, 591-598, 1993

A:Title: Purification and primary structure of ceratotoxin A and B, two antibacterial peptides
A:Reference number: A61613; MUID:93357786; PMID:8353519
A:Accession: B61613

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-29 <MAR>
C:Keywords: antibacterial; duplication

Query Match 50.0%; Score 25; DB 2; Length 29;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GSAFKALP 9
Db 3 GSAFKALP 11

RESULT 11

A10982
hypothetical protein STY4163 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhimurium
A>Note: this species has also been called Salmonella typhimurium
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: A10982
R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A:Reference number: A10982; PMID:11677608
A:Accession: A10982

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-43 <PAR>
A:Cross-references: GB:A1513382; PIDN:CAD07989.1; PID:g16504975; GSPDB:GN00176
C:Genetics:
A:Gene: STY4163

Query Match 50.0%; Score 25; DB 2; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFAKLQPR 9
Db 17 SRFAKALP 24

RESULT 12

B97874
degenerate transposase (orf2) [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: B97874

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Escherich, D.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; McRae, R.; LeBlanc, D.J.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; McRae, R.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5705-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B97874

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-49 <KOR>
A:Cross-references: GB:AE007317; PIDN:AAK98822.1; PID:g15457547; GSPDB:GN00174
C:Genetics:
A:Gene: IS1167-truncation

Query Match 50.0%; Score 25; DB 2; Length 49;
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSFAKL 7
 |||||
 Db 40 GSSFNKI 46

RESULT 13

S08565
 ribulose-bisphosphate carboxylase subunit-binding protein beta chain - garden pea (fragment)
 N:Alternate names: rubisco subunit-binding protein beta chain
 C:Species: Pisum sativum (garden pea)

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 12-Sep-1997

C:Accession: S08565

R:Musgrove, J.E.; Johnson, R.A.; Ellis, R.J.

Eur. J. Biochem. 163, 529-534, 1987

A:Title: Dissociation of the ribulosebiphosphate-carboxylase large-subunit binding protein
 A:Reference number: S07232; MUID:87161853; PMID:3549295

A:Accession: S08565

A:Molecule type: protein

A:Residues: 1-30 <MUS>

C:Comment: This protein binds the newly synthesized large subunit and the as newly imported
 C:Comment: This protein has ATPase activity.

C:Superfamily: chaperonin, groEL

C:Keywords: chloroplast; heterododecamer; molecular chaperone

Query Match 48.0%; Score 24; DB 2; Length 30;
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSFAKLQ 8
 |||||
 Db 10 GSAIRKLQ 17

RESULT 14

A47752

RNA recognition motif-type RNA-binding protein - fruit fly (Drosophila melanogaster) (fragment)

C:Species: Drosophila melanogaster

C>Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999

C:Accession: A47752

R:Kim, Y.J.; Baker, B.S.

Mol. Cell. Biol. 13, 174-183, 1993

A:Title: Isolation of RRM-type RNA-binding protein genes and the analysis of their relationship
 A:Reference number: A48110; MUID:93109300; PMID:8417324

A:Accession: A47752

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-39 <KIM>

A:Cross-references: GB:S51740; NID:g262273; PIDN:AAB24631.1; PID:g262274

A>Note: sequence extracted from NCBI backbone (NCBIN:121163, NCBIP:121164)

C:Genetics:

A:Gene: FlyBase:Rbp11

A:Cross-references: FlyBase:FBgn0010254

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein

Query Match 48.0%; Score 24; DB 2; Length 39;
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSPAKLQP 9
 |||||
 Db 16 SAFAKYGP 23

RESULT 15

PQ0617

DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 6 (fragment)

C:Species: human herpesvirus 6

C>Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 07-May-1999

C:Accession: PQ0617
 R:Ellinger, K.; Neipel, F.; Foa-Tomasi, L.; Campadelli-Piume, G.; Fleckenstein, B.
 J. Gen. Virol. 74, 495-500, 1993
 A:Title: The glycoprotein B homologue of human herpesvirus 6.
 A:Reference number: PQ0616; MUID:93187613; PMID:8383182
 A:Accession: PQ0617
 A:Molecule type: DNA
 A:Residues: 1-48 <ELL>
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA replication; nucleotidyltransferase

Query Match 48.0%; Score 24; DB 2; Length 48;
 Best Local Similarity 44.4%; Pred. No. 3.6e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSPAKLQPR 10
 |||||
 Db 22 SSVIRILPR 30

Search completed: January 10, 2003, 15:56:30
 Job time : 16.9091 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:16 ; Search time 9.09091 Seconds
(without alignments)
45.624 Million cell updates/sec

Title: C
Perfect score: 50
Sequence: 1 gsefaklqpr 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 3754

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	27	54.0	27 1 GRP_PIG	P01394 sus scrofa
2	26	52.0	43 1 PIV6_ADEB2	Q96627 bovine aden
3	25	50.0	11 1 Q20A_COMTE	P80464 comonomas t
4	25	50.0	25 1 GRP_SCYCA	P09472 scyliorhinu
5	25	50.0	29 1 CERB_CERCA	P36191 ceratilis c
6	23	46.0	37 1 V700_BORBU	O51643 borrelia bu
7	23	46.0	50 1 ISPI_GALME	P81505 galleria me
8	22.5	45.0	28 1 GRP_ALIMI	P31886 alligator m
9	22	44.0	47 1 V05_BPT7	P03777 bacterioph
10	21	42.0	7 1 LANG_CARUI	P36960 carnobacter
11	21	42.0	23 1 SODM_RANCA	P36215 rana catesb
12	21	42.0	26 1 RSHK_THERH	P32193 thermus the
13	21	42.0	27 1 GRP_CANFA	P08989 canis fam1
14	21	42.0	36 1 TXAM_METSE	P11495 metridium s
15	21	42.0	37 1 OP2B_OXYKI	P83349 oxyopes kit
16	21	42.0	45 1 TRPD_BACME	P70936 bacillus me
17	20	40.0	20 1 STYA_STYCL	P81469 styela clay
18	20	40.0	20 1 STVB_STYCL	P81470 styela clay
19	20	40.0	21 1 LPRM_CORDI	P21332 corynebacte
20	20	40.0	32 1 CRP_PLEPL	P12245 pleurocete
21	20	40.0	35 1 PYCI_SPIPL	P72506 spiliplina p
22	20	40.0	49 1 COXH_THUOB	P80977 thunnus obe
23	20	40.0	49 1 Y495_TREPA	O83508 treponema p
24	20	40.0	50 1 YF64_HABIN	P44256 haemophilus
25	19.5	39.0	24 1 FEED_ANNME	P80707 amycolatops
26	19	38.0	20 1 CATI_FASHS	Q09093 fasciola he
27	19	38.0	27 1 GRP_CHICK	P01395 gallus gall
28	19	38.0	29 1 DMD_RAT	P11530 rattus norv
29	19	38.0	39 1 Y410_BORBU	O51371 borrelia bu
30	19	38.0	42 1 MYXC_CRODU	P01778 crocalus du
31	19	38.0	43 1 MYX1_CROVC	P12028 crocalus vi
32	19	38.0	43 1 MYX2_CROVC	P12029 crocalus vi
33	19	38.0	43 1 MYXC_CROVH	P01477 crocalus vi

34	19	38.0	43 1 RR31_SPIOL	P47910 spinacia ol
35	19	38.0	45 1 MYX2_CROVH	P19861 crocalus vi
36	19	38.0	45 1 MYX_CROAD	P24330 crocalus ad
37	19	38.0	45 1 PH13_MYTCA	P11860 mytilus cal
38	19	38.0	45 1 VNSX_CVBM	P22052 bovine cal
39	19	38.0	46 1 GP54_BPSPI	O48408 bacterioph
40	19	38.0	47 1 FLGD_YEREN	Q56895 yerquina en
41	19	38.0	49 1 RK32_ASTLO	P24751 acetabula lon
42	18	36.0	10 1 SLAP_BACTG	P49325 bacillus th
43	18	36.0	12 1 TW2A_METWA	P80652 methanosarc
44	18	36.0	13 1 NP3_LYMST	P80180 lymanaea sca
45	18	36.0	15 1 F1BA_ANNAPL	P12801 anas platyr

ALIGNMENTS

RESULT 1	GRP_PIG	STANDARD:	PRT:	27 AA.
AC	P01394:			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Gastrin-releasing peptide (GRP) [Contains: Neuromedin C (GRP-10)].			
GN	GRP.			
OS	Sus scrofa (Pig), and			
OC	Cavia porcellus (Guinea pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
RN	NCBI_TaxID=9823, 10141;			
RP	SEQUENCE:			
RC	SPECTES=Pig;			
RX	MEDLINE=80042433; PubMed=496973;			
RA	McDonald T.J., Joernvall H., Nilsson G., Vagne M., Ghatei M.,			
RA	Bloom S.R., Mutt V.;			
RT	"Characterization of a gastrin releasing peptide from porcine			
RT	non-aneural gastric tissue.";			
RL	Biochem. Biophys. Res. Commun. 90:227-233(1979).			
RN	[2].			
RP	SEQUENCE OF 18-27.			
RX	SPECTES=Pig;			
RC	MEDLINE=84153890; PubMed=6546686;			
RA	Minamino N., Kangawa K., Matsuo H.;			
RT	"Neuromedin C: a bombesin-like peptide identified in porcine spinal			
RT	cord.";			
RL	Biochem. Biophys. Res. Commun. 119:14-20(1984).			
RN	[3].			
RP	SEQUENCE:			
RC	SPECTES=C.porcellus;			
RX	MEDLINE=8034977; PubMed=3668528;			
RA	Shaw C., Thim L., Conlon J.M.;			
RT	"Primary structure and tissue distribution of guinea pig gastrin-			
RT	releasing peptide.";			
RL	J. Neurochem. 49:1348-1354(1987).			
CC	- FUNCTION: GRP stimulates gastrin release as well as other			
CC	gastrointestinal hormones.			
CC	- SUBCELLULAR LOCATION: Secreted.			
CC	- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN			
CC	FAMILY.			
DR	PIR; A01562; RHPGA.			
DR	PIR; A60206; RHGSA.			
DR	InterPro; IPR000874; Bombesin.			
DR	Pfam; PF02044; Bombesin; 1.			
DR	PROSITE; PS00257; BOMBESIN; 1.			
KW	Bombesin family; Amidation.			
FT	PEPTIDE 18 27			NEUROMEDIN C.
FT	MOD_RES 27 27			AMIDATION.
FT	SEQUENCE 27 AA; 2806 MW; FA1017260210410F CRC64;			
SQ				

Query Match 54.0%; Score 27; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 23;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GSSFAKLQPR 10
| : | : | :
Db 8 GTVLAKMYPR 17

RESULT 2

PIV6_ADEB2
ID PIV6_ADEB2 STANDARD; PRT; 43 AA.
AC Q96627;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Minor capsid protein VI precursor (Fragment).
GN PVI.
OS Bovine adenovirus type 2 (Mastadenovirus bos2).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
ON NCBI_TaxID=114429;
RN [1]
RP SEQUENCE FROM N.A.

RA Kusvali M., Harrach B., Banrevi A., Evans P., Benko M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MINOR CAPSID PROTEIN THAT MAY ACT AS A LINK BETWEEN THE
CC EXTERNAL CAPSID AND THE INTERNAL DNA-PROTEIN CORE.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U44123; AAB16759.1; -.
DR InterPro; IPR004243; McpVI.
DR Pfam; PF02993; MCPVI; 1.
FT PROPEP 1 33 BY SIMILARITY
FT CHAIN 34 >43 MINOR CAPSID PROTEIN VI.
FT NON_TER 43 43
SQ SEQUENCE 43 AA; 4584 MW; 10F78E9678070306 CRC64;

Query Match 52.0%; Score 26; DB 1; Length 43;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GSSFAKLQPR 10
| : | : | :
Db 3 GINFSALAPR 12

RESULT 3

Q2OA_COMTE
ID Q2OA_COMTE STANDARD; PRT; 11 AA.
AC P80464;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinolone 2-oxido-reductase, alpha chain (EC 1.3.99.17) (Fragment).
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
ON NCBI_TaxID=285;
RN [1]
RP SEQUENCE.

RC STRAIN=63;
RX MEDLINE=96035889; PubMed=7556204;
RA Schach S., Teshisuka B., Fetzner S., Lingens F.;
RT "Quinolone 2-oxido-reductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT dioxxygenase from Comamonas testosteroni 63. The first two enzymes in
RT quinoline and 3-methylquinoline degradation.";
RL Eur. J. Biochem. 232:536-544 (1995).
CC -!- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)-2-OXO-
CC 1,2-DIHYDROQUINOLINE.

CC -!- CATALYTIC ACTIVITY: Quinolone + acceptor + H(2)O = isoquinolin-
CC 1(2H)-one + reduced acceptor.
CC -!- COFACTOR: FAD, MOLYBDENUM AND IRON-SULFUR.
CC -!- PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
CC (3-METHYL-)-QUINOLINE.
CC -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC TWO GAMMA CHAINS (PROBABLE).
KW Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1213 MW; 86909432281DC2CA CRC64;

Query Match 50.0%; Score 25; DB 1; Length 11;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSFAKLQPR 10
| : | : | :
Db 3 SDVAELKPR 11

RESULT 4

GRP_SCYCA
ID GRP_SCYCA STANDARD; PRT; 25 AA.
AC P09472;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gastrin-releasing peptide (GRP).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
ON NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.

RA MEDLINE=88137922; PubMed=3436516;
RX Conlon J.M., Henderson I.W., Thim L.;
RT "Gastrin-releasing peptide from the intestine of the elasmobranch
RT fish, Scyllorhinus canicula (common dogfish).";
RL Gen. Comp. Endocrinol. 68:415-420 (1987).
CC -!- FUNCTION: GRP stimulates gastrin release as well as other
CC gastrointestinal hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.

DR PIR; S06263; S06263.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT MOD_RES 25 25 AMIDATION.
SQ SEQUENCE 25 AA; 2781 MW; B735F911B89007F8 CRC64;

Query Match 50.0%; Score 25; DB 1; Length 25;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SFAKLQPR 10
| : | : | :
Db 8 SFPKMFPR 15

RESULT 5

CERB_CERCA
ID CERB_CERCA STANDARD; PRT; 29 AA.
AC P36191;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Ceratotoxin B.
GN CTXB.
OS Ceratitidis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Tephritidae; Tephritidae; Ceratitis.
 OC NCBI_TaxId=7213;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Female accessory gland;
 RX MEDLINE=93357786; PubMed=83353519;
 RA Marchini D., Giordano P.C., Amons R., Bernini L.F., Dallai R.;
 RT "Purification and primary structure of ceratotoxin A and B, two
 RT antibacterial peptides from the female reproductive accessory glands
 RT of the medfly *Ceratitis capitata* (Insecta:Diptera).";
 RL Insect Biochem. Mol. Biol. 23:591-598(1993).
 CC -1- FUNCTION: FEMALE-SPECIFIC PEPTIDES WITH POTENT ACTIVITY AGAINST
 CC GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA. THEY HAVE AS WELL
 CC HEMOLYTIC ACTIVITY. THESE PROTEINS ARE STABLE EVEN AT 100 DEGREES
 CC CELSIUS.
 CC -1- SUBUNIT: HOMOPOLYMER OF FOUR TO SIX SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: STRUCTURALLY RELATED TO CECROPINS, DEFENSINS AND
 CC APIDECINS.
 CC KW Insect immunity; Hemolysis; Antibiotic.
 SQ SEQUENCE 29 AA; 2861 MW; EE57F4EECB2DA6B0 CRC64;

Query Match 50.0%; Score 25; DB 1; Length 29;
 Best Local Similarity 55.6%; Pred. No. 67;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GSSFAKLP 9
 DB 3 GSAFKALP 11

RESULT 6

Y700 BORBU STANDARD; PRT; 37 AA.
 ID Y700 BORBU
 AC 051643;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BB0700.
 GN BB0700.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC NCBI_TaxId=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kertész A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utechtback T., Matthey L., McDonald L., Arlrich P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, *Borrelia*
 RT burgdorferi.";
 RL Nature 390:580-586(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----

DR EMBL AR001170; AAC67053.1; -
 DR TIGR BB0700; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 37 AA; 4267 MW; 39BAC907DE1B5B42 CRC64;

Query Match 46.0%; Score 23; DB 1; Length 37;
 Best Local Similarity 57.1%; Pred. No. 2.3e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GSSFAKL 7
 DB 20 GNNFGKL 26

RESULT 7

ISPL GALME STANDARD; PRT; 50 AA.
 ID ISPL GALME
 AC P81905;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Inducible serine protease inhibitor 1 (ISPL-1) (Fragment).
 DE Galleria mellonella (wax moth).
 OS Galleria mellonella
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC ditrysia; Pyralidae; Pyralidae; Galleriinae; Galleria.
 OC NCBI_TaxId=7137;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=20193629; PubMed=10727944;
 RA Froehls A.C., Kanost M.R., Goetz P., Vilcinskas A.;
 RT "Isolation and characterization of novel inducible serine protease
 RT inhibitors from larval hemolymph of the greater wax moth *Galleria*
 RT mellonella.";
 RL Eur. J. Biochem. 267:2046-2053(2000).
 CC -1- FUNCTION: INHIBITS TRYPSIN AND THE TOXIN PROTEASE PR2 OF M.
 CC ANISOPLAE. DOES NOT INHIBIT CHYMOTRYPSIN, SUBSTILISIN CARLSBERG,
 CC PROTEINASE K, PORCINE PANCREATIC ELASTASE AND THE TOXIN PROTEASE
 CC PR1 OF M. ANISOPLAE.
 CC -1- DEVELOPMENTAL STAGE: LAST INSTAR LARVAE.
 CC -1- INDUCTION: BY INHIBITOR.
 KW Serine protease inhibitor.
 FT NON_TER 50
 SQ SEQUENCE 50 AA; 5368 MW; AD67E6C9D8BC9254 CRC64;

Query Match 46.0%; Score 23; DB 1; Length 50;
 Best Local Similarity 44.4%; Pred. No. 3.2e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GSSFAKLP 9
 DB 5 GTFEXKNP 13

RESULT 8

GRP ALLMT STANDARD; PRT; 28 AA.
 ID GRP ALLMT
 AC P31886;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gastrin-releasing peptide (GRP) [Contains: Neuropeptide C (GRP-10)].
 OS Alligator mississippiensis (American alligator).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylia; Alligatorinae; Alligator.
 OC NCBI_TaxId=8496;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Stomach;
 RX MEDLINE=93324451; PubMed=8101369;
 RA Wang Y., Conlon J.M.;
 RT "Neuroendocrine peptides (NPY, GRP, VIP, somatostatin) from the brain
 RT and stomach of the alligator.";
 RL Peptides 14:573-579(1993).
 CC -1- FUNCTION: GRP stimulates gastrin release as well as other
 CC gastrointestinal hormones.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN FAMILY.

CC InterPro; IPR000874; Bombesin.

DR Pfam; PF02044; Bombesin; 1.

DR PROSITE; PS00257; BOMBESIN; 1.

KW Bombesin family; Amidation.

FT PEPTIDE 19 28 NEUROMEDIN C.

FT MOD RES 28 28 AMIDATION.

SQ SEQUENCE 28 AA; 2786 MW; A74BD0487D844963 CRC64;

Query Match 45.0%; Score 22.5; DB 1; Length 28;

Best Local Similarity 54.5%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GSS-PAKLOPR 10

Db 8 GSAPLAKTYPR 18

RESULT 9

V05_BPT7

ID -V05_BPT7 STANDARD; PRT; 47 AA.

AC P03777;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-MAR-1989 (Rel. 10, Last annotation update)

DE Gene 0.5 protein.

GN 0.5.

OS Bacteriophage T7.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;

OC T7-like viruses.

OX NCBI_TaxID=10760;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83241725; PubMed=6864790;

RA Dunn J.J., Studier F.W.;

RT "Complete nucleotide sequence of bacteriophage T7 DNA and the

locations of T7 genetic elements.";

RL J. Mol. Biol. 166:477-535 (1983).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=82078034; PubMed=7310871;

RA Dunn J.J., Studier F.W.;

RT "Nucleotide sequence from the genetic left end of bacteriophage T7

DNA to the beginning of gene 4";

RL J. Mol. Biol. 148:303-330 (1981).

CC -1- FUNCTION: THE FUNCTION OF THIS EARLY GENE PROTEIN IS UNKNOWN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

or send an email to license@sb-sib.ch).

CC -----

DR EMBL; V01146; CAA24386.1; -

DR EMBL; V01127; CAA24329.1; -

DR PIR; A04402; W08P57.

DR PIR; S42285; S42285.

SQ SEQUENCE 47 AA; 4745 MW; B07BC5B9FC12FA66 CRC64;

Query Match 44.0%; Score 22; DB 1; Length 47;

Best Local Similarity 66.7%; Pred. No. 4.9e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSPAK 6

Db 17 GASFGK 22

RESULT 10

LANC_CARUI

ID LANC_CARUI STANDARD; PRT; 7 AA.

AC P36960;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lantibiotic carnocin UI49 (Fragment).

OS Carnobacterium sp. (strain UI49).

OC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;

OC Carnobacterium.

OX NCBI_TaxID=35782;

RN [1]

RP SEQUENCE.

RX MEDLINE=92331768; PubMed=1622206;

RA Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,

RA Nes I.F.;

RT "Purification and characterization of a new bacteriocin isolated from

a Carnobacterium sp.";

RL Appl. Environ. Microbiol. 58:1417-1422 (1992).

CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).

CC ACTIVE ON GRAM-POSITIVE BACTERIA.

KW Antibiotic; Bacteriocin; Lantibiotic.

FT NON TER 7

SQ SEQUENCE 7 AA; 786 MW; 741776D05B05B810 CRC64;

Query Match 42.0%; Score 21; DB 1; Length 7;

Best Local Similarity 50.0%; Pred. No. 1.1e+05;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKLQPR 10

Db 2 SEIQPR 7

RESULT 11

SODM_RANCA STANDARD; PRT; 23 AA.

ID -SODM_RANCA

AC P362I5;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Superoxide dismutase [Mn], mitochondrial (EC 1.15.1.1) (Fragment).

OS Rana catesbeiana (Bull frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.

OX NCBI_TaxID=8400;

RN [1]

RP SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=87126854; PubMed=3492965;

RA Abe Y., Okazaki T.;

RT "Purification and properties of the manganese superoxide dismutase

from the liver of bullfrog, Rana catesbeiana.";

RL Arch. Biochem. Biophys. 253:241-248 (1987).

CC -1- FUNCTION: Destroys radicals which are normally produced within the

cells and which are toxic to biological systems.

CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -1- COFACTOR: Manganese.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

FAMILY.

DR InterPro; IPR001189; SODismutase.

DR Pfam; PF00081; sodfe, 1.

DR ProDom; PD000475; SODismutase; 1.

DR PROSITE; PS00088; SOD_MN; PARTIAL.

KW Oxidoreductase; Manganese; Mitochondrion.

FT NON TER 23

SQ SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04F625 CRC64;

Query Match 42.0%; Score 21; DB 1; Length 23;

Best Local Similarity 66.7%; Pred. No. 3.7e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      4 FAKLQ 9
Db      11 FGALQ 16

RESULT 12
RSHX THETH STANDARD; PRT; 26 AA.
AC P32153; P80383; Q9F2A8;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein Thx.
GN RPSU.
OS Thermus thermophilus, and
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxId=274, 271;
RN [1]
RP SEQUENCE.
RX MEDLINE=93363224; PubMed=8357533;
RA Choli T., Franceschi F., Yonath A., Wiltmann-Liebold B.;
RT "Isolation and characterization of a new ribosomal protein from the
RT thermophilic eubacteria, Thermus thermophilus, T. aquaticus and T.
RT flavus";
RL Biol. Chem. Hoppe-Seyler 374:377-383 (1993).
RN [2]
RP SEQUENCE.
RC SPECIES=T.thermophilus;
RX MEDLINE=95045586; PubMed=7957245;
RA Tsaloli P., Herfurth B., Choli T.;
RT "Purification and characterization of the 30S ribosomal proteins from
RT the bacterium Thermus thermophilus.";
RL Eur. J. Biochem. 226:169-177 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=T.thermophilus; STRAIN=HB / ATCC 27634;
RX MEDLINE=21421773; PubMed=11530930;
RA Leoniadou F., Triantafillidou D., Choli-Papadopoulos T.;
RT "On the characterization of the putative S20-thx operon of Thermus
RT thermophilus.";
RL Biol. Chem. 382:1001-1006 (2001).
CC -1- SIMILARITY: BELONGS TO THE S31E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ295159; CAC15068.1; -.
DR PIR; S33869; S33869.
KW Ribosomal protein.
FT INIT MET 0 0
SQ SEQUENCE 26 AA; 3206 MW; 8592641145D8C604 CRC64;

Query Match 42.0%; Score 21; DB 1; Length 26;
Best Local Similarity 37.5%; Pred. No. 4, 2e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      3 SFAKLQ 10
Db      16 TYGKYR 23

RESULT 13
GRP CANFA STANDARD; PRT; 27 AA.
AC P08989;
DT 01-NOV-1988 (Rel. 09, Created)

```

```

DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gastrin-releasing peptide (GRP) [Contains: Neuromedin C (GRP-10)].
GN GRP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxId=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=83213518; PubMed=6853532;
RA Reeve J.R. Jr., Walsh J.H., Chew P., Clark B., Hawke D.,
RA Shively J.E.;
RT "Amino acid sequences of three bombesin-like peptides from canine
RT intestine extracts";
RL J. Biol. Chem. 258:5582-5588 (1983).
CC -1- FUNCTION: GRP stimulates gastrin release as well as other
CC gastrointestinal hormones.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT PEPTIDE 18 27
FT MOD_RES 27 27
SQ SEQUENCE 27 AA; 2889 MW; 9D9317261B7C7D65 CRC64;

Query Match 42.0%; Score 21; DB 1; Length 27;
Best Local Similarity 40.0%; Pred. No. 4, 4e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy      1 GGSFAKLQ 10
Db      8 GTVLDRKY 17

RESULT 14
TXAM METSE STANDARD; PRT; 36 AA.
AC P1495;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metridin.
OS Metridium senile (Brown sea anemone) (Frieded sea anemone).
OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Actinoptaria;
OC Nymphaeaceae; Metridiidae; Metridium.
OX NCBI_TaxId=6115;
RN [1]
RP SEQUENCE.
RC TISSUE=Chidoblast;
RA Krebs H.C., Habermehl G.G.;
RT "Isolation and structural determination of a hemolytic active peptide
RT from the sea anemone Metridium senile.";
RL Naturwissenschaften 74:395-396 (1987).
CC -1- SUBCELLULAR LOCATION: Chidocyst and then secreted.
CC -1- SIMILARITY: BELONGS TO THE SEA ANEMONE POTASSIUM CHANNEL
CC INHIBITORY TOXIN FAMILY.
DR PIR; A27222; A27222.
DR InterPro; IPR003582; ShKT.
DR SMART; SM00254; ShKT; 1.
KW Toxin; Chidocyst.
SQ SEQUENCE 36 AA; 3974 MW; 5ED9CC73509E007F CRC64;

Query Match 42.0%; Score 21; DB 1; Length 36;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      3 SFAKLQ 8
Db      17 SFCKLE 22

```

```

RESULT 15
OP2B_OXYKI          STANDARD;          PRT;          37 AA.
ID   OP2B_OXYKI
AC   P83249;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Oxyopinin 2b (Oxx12b).
OS   Oxyopes kitabensis (Wolf spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Oxyopidae; Oxyopes.
OX   NCBI_taxID=184771;
[1]
RN   SEQUENCE, FUNCTION, TISSUE SPECIFICITY, SUBCELLULAR LOCATION, MASS
RP   SPECTROMETRY, AND CIRCULAR DICHROISM ANALYSIS.
RC   TISSUE=Venom;
RX   PubMed=11976325;
RA   Corzo G., Villegas E., Gomez-Lagunas F., Possani L.D.,
RA   Belokoneva O.S., Nakajima T.;
RT   "Oxyopinins, large amphipathic peptides isolated from the venom of the
RT   wolf spider Oxyopes kitabensis with cytolytic properties and positive
RT   insecticidal cooperativity with spider neurotoxins.";
RL   J. Biol. Chem. 277:23627-23637(2002).
CC   -!- FUNCTION: Disrupts biological membranes, particularly those rich
CC   in phosphocholine. Has antimicrobial activity against Gram-
CC   negative bacterium E.coli, Gram-positive bacteria B.subtilis and
CC   S.aureus, and hemolytic activity against sheep, pig and guinea pig
CC   red blood cells. Has insecticidal activity against S.frugiperda
CC   ovarian cells by opening non-selective ion channels. Enhances the
CC   insecticidal activity of spider venom neurotoxic peptides.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MASS SPECTROMETRY: MW=4146.9; METHOD=MALDI.
KW   Antibiotic; Toxin; Hemolysis; Cytolysis.
SQ   SEQUENCE 37 AA; 4147 MW; 590B2ED8BE99A1EB CRC64;

Query Match          42.0%; Score 21; DB 1; Length 37;
Best Local Similarity 66.7%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSFAKL 7
    | | | |
Db 4 SGFAKI 9

Search completed: January 10, 2003, 15:55:49
Job time : 11.0909 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 15:55:17 ; Search time 29.5455 Seconds
(without alignments)
69.739 Million cell updates/sec

Title: C
Perfect score: 50
Sequence: 1 gssfaklqpr 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 33835

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_ricent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeophages: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	32	64.0	43	12	Q65289
2	28	56.0	16	4	Q16350
3	28	56.0	19	10	O80997
4	27	54.0	37	16	O9HY43
5	27	54.0	42	2	P77063
6	27	54.0	43	2	P77064
7	27	54.0	43	2	P77066
8	27	54.0	43	2	P77498
9	27	54.0	43	2	P77514
10	26	52.0	39	5	Q21886
11	26	52.0	39	11	O9R137
12	25	50.0	29	15	O89348
13	25	50.0	33	11	O88440
14	25	50.0	42	12	O91X54
15	25	50.0	43	16	O822A2
16	25	50.0	43	16	O822A2

17	25	50.0	44	4	Q96CFL	Q96CFL homo sapien
18	25	50.0	45	13	Q9PRX6	Q9PRX6 xenopus lae
19	24.5	49.0	27	12	O56530	O56530 hepatitis c
20	24	48.0	26	11	O8VIN3	O8VIN3 mus musculus
21	24	48.0	35	11	O8VB06	O8VB06 mus musculus
22	24	48.0	39	5	Q26280	Q26280 drosophila
23	24	48.0	39	11	O8VBX8	O8VBX8 mus musculus
24	24	48.0	41	11	O8VB03	O8VB03 mus musculus
25	24	48.0	48	11	O8VIN5	O8VIN5 mus musculus
26	24	48.0	50	16	O9KTR0	O9KTR0 vibrio chol
27	23	46.0	13	13	O8UHO0	O8UHO0 gallus gall
28	23	46.0	18	6	O95MB1	O95MB1 equus caball
29	23	46.0	26	13	O8UVE1	O8UVE1 gallus gall
30	23	46.0	27	11	O9CSC1	O9CSC1 mus musculus
31	23	46.0	30	16	O9K239	O9K239 chlamydia p
32	23	46.0	39	16	O9J098	O9J098 neisseria m
33	23	46.0	40	16	O8X2K7	O8X2K7 escherichia
34	23	46.0	43	4	O9H480	O9H480 homo sapien
35	23	46.0	43	6	Q28816	Q28816 phoca vitul
36	23	46.0	43	6	Q28444	Q28444 hydurga le
37	23	46.0	43	6	Q28458	Q28458 mirounga an
38	23	46.0	43	6	Q27942	Q27942 arctocepal
39	23	46.0	43	6	Q28456	Q28456 leprocyctot
40	23	46.0	44	2	Q44544	Q44544 anabaena va
41	23	46.0	45	16	O9JSM4	O9JSM4 neisseria m
42	23	46.0	46	12	O991B9	O991B9 hepatitis c
43	23	46.0	50	6	Q28491	Q28491 mirounga le
44	23	46.0	50	16	O51559	O51559 borrelia bu
45	22	44.0	15	10	O8W4X5	O8W4X5 arabidopsis

ALIGNMENTS

RESULT 1
Q65289 ID Q65289 PRELIMINARY; PRT; 43 AA.
AC Q65289; ID Q65289; PRT; 43 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-UN-2002 (TREMBLrel. 21, Last annotation update)
DE Protein 3a (Fragment).
OS Human adenovirus type 3.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=45659;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE 3;
RX MEDLINE=94357446; PubMed=8076828;
RA Cuzange A., Chroboczek J., Jacot B.;
RT "The penton base of human adenovirus type 3 has the RGD motif."
RL Gene 146:257-259 (1994).
DR EMBL; Z29487; CAA82621.1; -
DR InterPro; IPR003479; Hex_I1a.
DR Pfam; PF02455; Hex_I1a; 1.
FT NON_TER 1 1
SQ SEQUENCE 43 AA; 4754 MW; F8E42D960369585A CRC64;
Query Match Score 32; DB 12; Length 43;
Best Local Similarity 60.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GSSFALQPR 10
DB 29 GNFALRPR 38
RESULT 2
Q16350 ID Q16350 PRELIMINARY; PRT; 16 AA.
AC Q16350; ID Q16350; PRT; 16 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 Cystathionine beta-synthase (Fragment).
 Homo sapiens (Human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 [1]
 NCBI_TaxID=9606;
 SEQUENCE FROM N.A.
 MEDLINE=95282779; PubMed=7762555;
 Sebastian G., Sperandeo M.P., Panico M., de Franchis R., Kraus J.P.,
 Andria G.;
 "The molecular basis of homocystinuria due to cystathionine beta-
 synthase deficiency in Italian families, and report of four novel
 mutations";
 Am. J. Hum. Genet. 56:1324-1333(1995).
 DR EMBL; S78267; AAB34404.1; -;
 FT NON TER 1
 SQ SEQUENCE 16 AA; 1538 MW; 07455C146B001686 CRC64;
 Query Match 56.0%; Score 28; DB 4; Length 16;
 Best Local Similarity 55.6%; Pred. No. 46;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GSSFAKLQ 9
 Db 8 GGAFAGLEP 16
 RESULT 3
 ID O80997 PRELIMINARY; PRT; 19 AA.
 AC O80997;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 GN AT225390 protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 Salzberg S.L., Fraser C.M., Venter J.C.;
 "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 Nature 402:761-768(1999).
 RL Nature 402:761-768(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004747; AAC31236.1; -;
 SQ SEQUENCE 19 AA; 2134 MW; A4D7C1507A25FFA7 CRC64;
 Query Match 56.0%; Score 28; DB 10; Length 19;
 Best Local Similarity 60.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GSSFAKLQ 10
 Db 3 GSQREKLQ 12

RESULT 4
 ID Q9HY43 PRELIMINARY; PRT; 37 AA.
 AC Q9HY43;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
 DE Hypothetical protein PA3577.
 GN PA3577.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 Brody L.B., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 opportunistic pathogen";
 Nature 406:959-964(2000).
 RL Nature 406:959-964(2000).
 RW EMBL; AB004778; AAG06965.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 37 AA; 4372 MW; 0C68B97A779C241E CRC64;
 Query Match 54.0%; Score 27; DB 16; Length 37;
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GSSFAKLQ 8
 Db 7 GESFGLQ 14
 RESULT 5
 ID P77063 PRELIMINARY; PRT; 42 AA.
 AC P77063;
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Variant shiga-like toxin II VT subunit A (Fragment).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
 J. Clin. Microbiol. 0:0-0(1995).
 RL EMBL; U41250; AAB40560.1; -;
 DR HSP; P08027; 1BQV.
 DR InterPro; IPR003189; SLT_beta.
 DR Pfam; PF02258; SLT_beta; 1.
 FT NON TER 1
 SQ SEQUENCE 42 AA; 4576 MW; BC7DC6B741648385 CRC64;
 Query Match 54.0%; Score 27; DB 2; Length 42;
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GSSFAKLQ 8
 Db 31 GSGFAEVQ 38
 RESULT 6
 ID P77064

```

ID P77064 PRELIMINARY; PRT; 43 AA.
AC P77064;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Variant shiga-like toxin II VT subunit A (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
RL J. Clin. Microbiol. 0:0-0(1995).
DR EMBL; U41251; AAB40561.1; -.
DR HSSP; P08027; IBOV.
DR InterPro; IPR003189; SLT_beta.
DR Pfam; PF02258; SLT_beta; 1.
FT NON_TER 1
SQ SEQUENCE 43 AA; 4822 MW; 65792D9896947E0C CRC64;

Query Match 54.0%; Score 27; DB 2; Length 43;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSFAKQ 8
Db 32 GSGFAEVQ 39

RESULT 7
P77066 PRELIMINARY; PRT; 43 AA.
ID P77066;
AC P77066;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Variant shiga-like toxin II VT subunit A (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
RL J. Clin. Microbiol. 0:0-0(1995).
DR EMBL; U41250; AAB40570.1; -.
DR HSSP; P08027; IBOV.
DR InterPro; IPR003189; SLT_beta.
DR Pfam; PF02258; SLT_beta; 1.
FT NON_TER 1
SQ SEQUENCE 43 AA; 4354 MW; 71CD1CE37212ED34 CRC64;

Query Match 54.0%; Score 27; DB 2; Length 43;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSFAKQ 8
Db 32 GSGFAEVQ 39

RESULT 8
P77498 PRELIMINARY; PRT; 43 AA.
ID P77498;
AC P77498;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Variant SHIGA-like toxin II VT subunit A (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.

```

```

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41255; AAB40565.1; -.
DR EMBL; U41252; AAB40562.1; -.
DR HSSP; P08027; IBOV.
DR InterPro; IPR003189; SLT_beta.
DR Pfam; PF02258; SLT_beta; 1.
FT NON_TER 1
SQ SEQUENCE 43 AA; 4739 MW; BC7D8C2938C9385 CRC64;

Query Match 54.0%; Score 27; DB 2; Length 43;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSFAKQ 8
Db 32 GSGFAEVQ 39

RESULT 9
P77514 PRELIMINARY; PRT; 43 AA.
ID P77514;
AC P77514;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Variant SHIGA-like toxin II VT subunit A (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41258; AAB40568.1; -.
DR EMBL; U41256; AAB40566.1; -.
DR HSSP; P08027; IBOV.
DR InterPro; IPR003189; SLT_beta.
DR Pfam; PF02258; SLT_beta; 1.
FT NON_TER 1
SQ SEQUENCE 43 AA; 4797 MW; BC7D8C2938C9385 CRC64;

Query Match 54.0%; Score 27; DB 2; Length 43;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSFAKQ 8
Db 32 GSGFAEVQ 39

RESULT 10
P77065 PRELIMINARY; PRT; 44 AA.
ID P77065;
AC P77065;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Variant shiga-like toxin II VT subunit A (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin I.E., Bacon D.J., Tyler S.D., Munro C.K., Johnson W.M.;
RL J. Clin. Microbiol. 0:0-0(1995).
DR EMBL; U41253; AAB40563.1; -.
DR HSSP; P08027; IBOV.

```

```

DR InterPro: IPR003189; SLT beta.
PFam: PF02258; SLT_beta; 1.
FT NON_TER 1
SQ SEQUENCE 44 AA; 4851 MW; 6579398F654F550C CRC64;

Query Match 54.0%; Score 27; DB 2; Length 44;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSFAKLQ 8
   |||::|
Db 35 GSGFAEVQ 42

RESULT 11
Q21886 Q21886 PRELIMINARY; PRT; 39 AA.
AC Q21886;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 4.6 kDa protein.
GN R09H3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Martin J.;
RT "The sequence of C. elegans cosmid R09H3."
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RW EMBL; U58740; AAB00612.1; -.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4576 MW; 524E24643534359B CRC64;

Query Match 52.0%; Score 26; DB 5; Length 39;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AKLQPR 10
   |||
Db 2 AKQPR 7

RESULT 12
Q9R137 Q9R137 PRELIMINARY; PRT; 39 AA.
AC Q9R137;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cyclophilin A (Fragment).
GN PFIA OR CYPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=20422670; PubMed=10964515;
RA Colgan J., Asmal M., Luban J.;
RT "Isolation, characterization and targeted disruption of mouse Ppia:
RT cyclophilin A is not essential for mammalian cell viability."
RL Genomics 68:167-178(2000).
DR EMBL; AF171073; AAD50996.1; -.
DR HSSP; P05092; 2CPL.
DR MGD; MGI:97749; Ppia.
DR InterPro; IPR002130; CSA_PPIase.
DR Pfam; PF0160; pro_isomerase; 1.
DR PROSITE; PS50072; CSA_PPIASE_2; 1.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4324 MW; CB53F70E1092889C CRC64;

Query Match 52.0%; Score 26; DB 11; Length 39;
Best Local Similarity 55.6%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSFAKLQ 9
   |||::|
Db 26 GGSFHRIP 34

RESULT 13
O89348 O89348 PRELIMINARY; PRT; 29 AA.
AC O89348;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98282291; PubMed=9616227;
RA Brander C., Hartman K.E., Trocha A.K., Jones N.G., Johnson P.R.,
RA Korber B., Wentworth P., Buchbinder S.P., Wolinsky S., Walker B.D.,
RA Kalams S.A.;
RT "Lack of strong immune selection pressure by the immunodominant, HLA-
RT A*0201-restricted cytotoxic T lymphocyte response in chronic human
RT immunodeficiency virus-1 infection."
RL J. Clin. Invest. 101:2559-2566(1998).
DR EMBL; AF017816; AAC29148.2; -.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 29
SQ SEQUENCE 29 AA; 3192 MW; 5B8AF6E47A3FD746 CRC64;

Query Match 50.0%; Score 25; DB 15; Length 29;
Best Local Similarity 55.6%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GSSFAKLQ 9
   |||
Db 2 GKXFELLQ 10

RESULT 14
O88440 O88440 PRELIMINARY; PRT; 33 AA.
ID O88440
AC O88440;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```


DE Nonmuscle tropomyosin 5 (Fragment).
 GN TPM5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxId=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE=98325069; PubMed=9660825;
 RA Dufour C., Weinberger R.P., Schvezov G., Jeffrey P.L., Gunning P.;
 RT "Splicing of two internal and four carboxyl-terminal alternative exons
 in nonmuscle tropomyosin 5 pre-mRNA is independently regulated during
 development.";
 RL J. Biol. Chem. 273:18547-18555(1998).
 DR EMBL; AF053360; AAC27291.1; -;
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00261; Tropomyosin; 1.
 FT NON_TER
 SQ SEQUENCE 33 AA; 3723 MW; 5FF5861B89A38F48 CRC64;

Query Match 50.0%; Score 25; DB 11; Length 33;
 Best Local Similarity 62.5%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SSFAXLQP 9
 Db 25 TSFASLQP 32

RESULT 15

O9IX54
 ID O9IX54 PRELIMINARY; PRT; 42 AA.
 AC O9IX54;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Major core protein (Fragment).
 OS Possum adenovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Atadenovirus.
 OK NCBI_TaxId=121816;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Thomson D.M., Meers J.;
 RT "Molecular confirmation of an adenovirus in brushtail possums
 (Trichosurus vulpecula).";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF249333; AAF65557.1; -;
 DR InterPro; IPR004912; Adeno VII.
 DR Pfam; PF03228; Adeno VII; 1.
 FT NON_TER
 SQ SEQUENCE 42 AA; 4765 MW; 87C1D4978D1D13EC CRC64;

Query Match 50.0%; Score 25; DB 12; Length 42;
 Best Local Similarity 55.6%; Pred. No. 5.6e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GSSFAKLQP 9
 Db 25 GLRFSKRP 33

Search completed: January 10, 2003, 15:57:43
 Job time : 31.5455 secs

THIS PAGE BLANK (US70)